

Access DB# 92579

# SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

**If more than one search is submitted, please prioritize searches in order of need.**

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

**BEST AVAILABLE COPY**

Jan Delaval  
Reference Librarian  
Biotechnology & Chemical Library  
CM1 1E07 - 703-308-4498  
jan.delaval@uspto.gov

\*\*\*\*\*

## STAFF USE ONLY

|  | Type of Search           | Vendors and cost where applicable |
|--|--------------------------|-----------------------------------|
| Searcher: <u>Jan</u>                   | NA Sequence (#) _____    | STN _____                         |
| Searcher Phone #: <u>4498</u>          | AA Sequence (#) <u>✓</u> | Dialog _____                      |
| Searcher Location: _____               | Structure (#) _____      | Questel/Orbit _____               |
| Date Searcher Picked Up: <u>5/1/03</u> | Bibliographic _____      | Dr.Link _____                     |
| Date Completed: <u>5/1/03</u>          | Litigation _____         | Lexis/Nexis _____                 |
| Searcher Prep & Review Time: _____     | Fulltext _____           | Sequence Systems <u>✓</u>         |
| Clerical Prep Time: <u>20</u>          | Patent Family _____      | WWW/Internet _____                |
| Online Time: <u>15</u>                 | Other _____              | Other (specify) _____             |

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:31:42 ; Search time 27.2873 Seconds  
(without alignments)  
1602.986 Million cell updates/sec

Title: US-09-757-982-5  
Perfect score: 2375  
Sequence: 1 MSSLGASFVQIKFDDLQFFE.....GDDDDDDDGEEEDNDMDNSE 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
al number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID     | Description        |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1          | 601   | 25.3        | 954    | 1  | S68178 | mixed-lineage prot |
| 2          | 593   | 25.0        | 394    | 2  | JU0229 | mixed-lineage prot |
| 3          | 581.5 | 24.5        | 847    | 1  | A53800 | mixed-lineage prot |
| 4          | 565.5 | 23.8        | 668    | 2  | JC2363 | protein kinase (EC |
| 5          | 565.5 | 23.8        | 888    | 2  | A55318 | serine/threonine p |
| 6          | 560.5 | 23.6        | 406    | 2  | T52626 | probable mitogen-a |
| 7          | 560.5 | 23.6        | 888    | 2  | JC5399 | dual leucine zippe |
| 8          | 545.5 | 23.0        | 328    | 2  | T16747 | hypothetical prote |
| 9          | 529   | 22.3        | 886    | 2  | T48544 | MAP3K delta-1 prot |
| 10         | 510.5 | 21.5        | 1030   | 2  | F96763 | hypothetical prote |
| 11         | 505   | 21.3        | 848    | 2  | B87950 | protein F33E2.2 li |
| 12         | 505   | 21.3        | 855    | 2  | T20082 | hypothetical prote |
| 13         | 495   | 20.8        | 982    | 2  | T06576 | probable protein k |
| 14         | 488   | 20.5        | 821    | 2  | T48400 | serine/threonine-p |
| 15         | 483   | 20.3        | 579    | 2  | JC5955 | transforming growt |
| 16         | 480   | 20.2        | 829    | 2  | T07406 | probable protein k |
| 17         | 478.5 | 20.1        | 1015   | 2  | T10671 | probable serine/th |
| 18         | 475.5 | 20.0        | 412    | 2  | T10671 | protein kinase hom |
| 19         | 471.5 | 19.9        | 567    | 2  | JC5957 | transforming growt |
| 20         | 471.5 | 19.9        | 606    | 2  | JC5956 | transforming growt |
| 21         | 470   | 19.8        | 736    | 2  | T05137 | protein kinase hom |
| 22         | 468   | 19.7        | 390    | 2  | T01451 | protein kinase hom |
| 23         | 466   | 19.6        | 462    | 2  | S29851 | protein kinase 6 ( |
| 24         | 460   | 19.4        | 963    | 2  | T09911 | probable serine/th |
| 25         | 454.5 | 19.1        | 553    | 2  | T04683 | hypothetical prote |
| 26         | 447.5 | 18.8        | 407    | 2  | G84635 | probable protein k |
| 27         | 445.5 | 18.8        | 546    | 2  | D84555 | probable protein k |
| 28         | 441   | 18.6        | 475    | 2  | T12955 | probable protein k |
| 29         | 434   | 18.3        | 981    | 1  | FOXVGM | gag-abl polyprotei |

|    |       |      |      |   |        |                    |
|----|-------|------|------|---|--------|--------------------|
| 30 | 424   | 17.9 | 1257 | 2 | T00486 | serine/threonine-s |
| 31 | 421.5 | 17.7 | 738  | 2 | F96701 | hypothetical prote |
| 32 | 410   | 17.3 | 1130 | 1 | TVHUA  | protein-tyrosine k |
| 33 | 407   | 17.1 | 842  | 2 | T32258 | hypothetical prote |
| 34 | 405.5 | 17.1 | 697  | 2 | A26132 | gag-abl-pol polyp  |
| 35 | 405   | 17.1 | 1123 | 2 | A39962 | kinase-related tra |
| 36 | 405   | 17.1 | 1146 | 2 | B35962 | protein-tyrosine k |
| 37 | 405   | 17.1 | 1182 | 2 | A35962 | protein-tyrosine k |
| 38 | 404   | 17.0 | 988  | 2 | F86316 | protein T10022.13  |
| 39 | 403.5 | 17.0 | 1147 | 2 | F86297 | hypothetical prote |
| 40 | 402   | 16.9 | 1520 | 1 | TVFPA  | protein-tyrosine k |
| 41 | 401.5 | 16.9 | 1584 | 2 | T18276 | protein-tyrosine k |
| 42 | 400.5 | 16.9 | 1094 | 2 | S49313 | protein kinase - s |
| 43 | 400   | 16.8 | 1248 | 2 | B96827 | hypothetical prote |
| 44 | 398.5 | 16.8 | 1171 | 2 | T12956 | hypothetical prote |
| 45 | 396.5 | 16.7 | 545  | 2 | T05675 | hypothetical prote |

ALIGNMENTS

RESULT 1

S68178  
mixed-lineage protein kinase 2 (EC 2.7.1.1) - human  
C/Species: Homo sapiens (man)  
C/Date: 28-Oct-1996 #sequence revision 13-Mar-1997 #text\_change 11-Jun-1999  
C/Accession: S68178; I38044; S32468  
R/Dorow, D.S.; Devereux, L.; Tu, G.F.; Price, G.; Nicholl, J.K.; Sutherland, G.R.; Simp.  
Eur. J. Biochem. 234, 492-500, 1995  
A/Title: Complete nucleotide sequence, expression, and chromosomal localisation of huma  
A/Reference number: S68178; MUID:96128179; PMID:8536694  
A/Accession: S68178  
A/Molecule type: mRNA  
A/Residues: 1-954 <DOR>  
A/Cross-references: EMBL:X08046; NID:g971419; PIDN:CAA62351.1; PID:g971420  
R/Katoch, M.; Hirai, M.; Sugimura, T.; Terada, M.  
Oncogene 10, 1447-1451, 1995  
A/Title: Cloning and characterization of MST, a novel (putative) serine/threonine kinas  
A/Reference number: I38044; MUID:95249256; PMID:7731697  
A/Accession: I38044  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-461, 'A', 'V', '465-470, 'S', '472-806, 'R', '808-817, 'A', '819-954 <RES>  
A/Cross-references: EMBL:Z48615; NID:g758592; PIDN:CAA88531.1; PID:g758593  
R/Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.  
Eur. J. Biochem. 213, 701-710, 1993  
A/Title: Identification of a new family of human epithelial protein kinases containing  
A/Reference number: S32467; MUID:93238756; PMID:8477742  
A/Accession: S32468  
A/Molecule type: mRNA  
A/Residues: 244-464, 'AQAAGRQPHQPALWL' <DO2>  
C/Genetics:  
A/Gene: GDB:MLK2; GDB:MST  
A/Cross-references: GDB:362654; GDB:624810; OMIM:600137  
A/Map position: 19q13.1-19q13.2  
C/Superfamily: mixed-lineage protein kinase 2; protein kinase homology; SH3 homology  
C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein  
F;23-76/Domain: SH3 homology <SH3>  
F;96-364/Domain: protein kinase homology <KIN>  
F;104-112/Region: protein kinase ATP-binding motif  
F;384-405/Region: leucine zipper motif  
F;419-440/Region: leucine zipper motif  
F;449-463/Region: basic  
F;125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match

Best Local Similarity 25.3%; Score 601; DB 1; Length 954;  
Matches 155; Conservative 63; Mismatches 160; Indels 60; Gaps 10;

OY 10 QIKFDDLQFFENCGGSGFSVYRAKWISQDKEVAVKL-LKIEK-----EAIL 57  
Db 92 EIPFHELQLEIITGVGFGKYRYALW--RGEVAVYXARLDPEKDPAVTAEQVCGEARLF 149

[illegible]

## RESULT 2

mixed-lineage protein kinase 1 - human  
C/Species: Homo sapiens (man)  
C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 17-Mar-1999  
C/Accession: S332467; J00229  
R/Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.  
Eur. J. Biochem. 213, 701-710, 1993  
A/Title: Identification of a new family of human epithelial protein kinases containing  
A/Reference number: S332467; MUID:93238756; PMID:8477742  
A/Accession: S332467  
A/Molecule type: mRNA  
A/Residues: 1-394 <DO2>  
C/Genetics:  
A/Gene: GDB:MLKL  
A/Cross-references: GDB:141921; OMIM:600136  
A/Map position: 14q24.3-14q31  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein kinase  
A/1-269/Domain: protein kinase homology <KIN>  
A/1-268/Domain: catalytic <CAT>  
E/9-17/Region: protein kinase ATP-binding motif  
F/289-310/Region: leucine zipper motif  
F/324-345/Region: leucine zipper motif  
F/354-368/Region: basic

|                           |       |                    |            |             |
|---------------------------|-------|--------------------|------------|-------------|
| Query Match               | 25.0% | Score 593;         | DB 2;      | Length 394; |
| Best Local Similarity     | 39.3% | Pred. No. 7.1e-25; |            |             |
| Matches 133; Conservative | 59;   | Mismatches 114;    | Indels 32; | Gaps 5;     |

|    |     |                           |                    |               |             |             |
|----|-----|---------------------------|--------------------|---------------|-------------|-------------|
| QY | 15  | DLOFFENCGSGSGSVYRAKWI     | SQDKEVAVK-----     | KLKIKEAEILSVL | SH          | 62          |
|    |     | :   :   :   :   :   :     | :   :   :   :   :  | : : : : : : : |             |             |
| Db | 2   | ELTLEEIIIGIGFGKYRAFWIGD-- | EVAVKAAARHDPDEDISQ | TENVRQ        | EAKL        | FAMLRKH 59  |
| QY | 63  | RNIIOFYGVILEPBNYIVTEYASL  | GSLYDYINSNRSEEMDM  | DMHTWATD      | VAKGMHYH    | 122         |
|    |     | :   :   :   :   :   :     | : : :   : : :      | : : :   : : : |             |             |
| Db | 60  | PNIIALRGVCLKEPNLCLVMEF    | ARGPLNRVYLSGKR--   | IPDILVNWAQ    | IARGMNYH    | 116         |
| QY | 123 | MEAPVKVIHRDLKSRNVV-----   | AADGVLCICDFGASR    | FRHNHTHMS     | LVGTFPMMA   | 174         |
|    |     | :   :   :   :   :   :     | : : :   : : :      | : : :   : : : |             |             |
| Db | 117 | DEAIVPIIHRDLKSSNII        | LQKVENGDLSNKILKITD | FLAREWHRTTK   | MSAAGTYANMA | 176         |
| QY | 175 | PAVIOQLPVSFETCDTYSYGV     | LWEMLTREYDFKGLGLQ  | VAWLVEKNER    | LTI         | PSSCPRS 234 |
|    |     | :   :   :   :   :   :     | : : :   : : :      | : : :   : : : |             |             |

```

Db      177 PEVIRASMFSGSDVWSYGVLWELLTGEVPPFRGIDGLRVAAYGVAMNKALPIPTCPEP 236
QY      235 FAEILLHQCWEADAKRRPSFKQIITLESMSNDTSLPDKCNSFLHNKAEMRCIEATLERL 294
        ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      237 FAKIMEDCWNPDPHSRPSFTNILDQLTTIEBSGFPEMPKDSFHCLQDNWKHEIQEMFDQL 296
QY      295 KKLIERDLSFKEOEL-----KERERRLKWMEQKLTQ 325
        ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      297 RAKEKELRTEWEEELTRALQOKNOEELLRRRQELAE 334

```

### RESULT 3

mixed-lineage protein kinase (EC 2.7.1.1-) 3 - human  
N/Alternate names: protein kinase PTK1; protein kinase SPRK  
C/Species: Homo sapiens (man)  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C/Accession: A53800; 158395  
R/Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.  
J. Biol. Chem. 269, 15092-15100, 1994  
A/Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont  
A/Reference number: A53800; MUID:94253068; PMID:8195146  
A/Accession: A53800  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-847 <GAL>  
A/Cross-references: GB:U07747; NID:G464027; PIDN:AAA19647.1; PID:G464028  
R/Ing, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lassam, N.J.  
Oncogene 9, 1745-1750, 1994  
A/Title: MLK-3: Identification of a widely-expressed protein kinase bearing an SH3 domain  
A/Reference number: 158395; MUID:94239754; PMID:8183572  
A/Accession: 158395  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-847 <RES>  
A/Cross-references: GB:L32976; NID:G488295; PIDN:AAA59859.1; PID:G488296  
C/Genetics:  
A/Gene: GDB:MLK3; PTK1, SPRK  
A/Cross-references: GDB:134755; OMIM:600050  
A/Map position: 11q13.1-11q13.3  
C/Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology  
C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein  
F/48-100/Domain: SH3 homology <SH32>  
F/115-383/Domain: protein kinase homology <KIN>  
F/123-131/Region: protein kinase ATP-binding motif  
F/403-424/Region: leucine zipper motif  
F/438-459/Region: leucine zipper motif  
F/468-482/Region: basic

|                       |                  |                    |            |             |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match           | 24.5%;           | Score 581.5;       | DB 1;      | Length 847; |
| Best Local Similarity | 37.4%;           | Pred. No. 6.8e-24; |            |             |
| Matches 132;          | Conservative 60; | Mismatches 122;    | Indels 39; | Gaps 7;     |

|    |     |                       |   |  |     |
|----|-----|-----------------------|---|--|-----|
| OY | 13  | FDDLQFFENC            | GGSFGSVYRAK   | WISQ-----DKEVAVKLLIKIEKAELSVLS               | 61  |
| Dd | 114 | FQELRL                | EVEVIGIGF   | GKVYRGSMGELVAVKAARQDPEDISV-TAESVRQEARLFAMLA  | 172 |
| OY | 62  | HNNIIQFYGVILEP        | PNYGIAT   | EYASLGLDYIINSNRSEEMDMDHIM-TWATDVAKGMHY       | 120 |
| Dd | 173 | HNNIIA                | LKAVCL  | EEPNLCUWEYAAGPLSRALAGRVP---PHVLWNVAVQIARGMHY | 228 |
| OY | 121 | LHMEAPVKVIHRDLKS      | RNVV----IAADGV---- <td>CFGASRFHNHTTHMSLVGTFPW</td> <td>172</td> | CFGASRFHNHTTHMSLVGTFPW                       | 172 |
| Dd | 229 | LHCEALPVPIHRDLKS      | NNILLLOPIESDDMEHKTLKITDFGLAREMKTQTOMSAAGTYAW                    | 288  |     |
| OY | 173 | MAREVIOQLPVSETCDTYSYG | VULWEMLTREVPFKGLEGLQVAMLVVEKNERLTIIPSSCP                        | 232  |     |
| Dd | 289 | MAREVIKASTFSKGS       | DVWSFCGVILWELLITGEVPRYGRIGIDCLAVALGYAVNKLTLPIPTSTCP             | 348  |     |
| OY | 233 | RSFAELLHQWEADAKRPS    | FKOII   | SILBSMSNDTSLPDKCNSFLHNKAEMRCIEATLE           | 292 |
| Dd | 349 | EPFAQLMADCWAQDPHRR    | RPDAFSTILOOLEALEAQVLR   | EMPRDSFHSMOEGWKREIQGLFD                      | 408 |

QY 293,RLKLERDLSFKEQE-----LKERERRLKMEQKLTQSNTPLL 331  
Db 409 ELRAKEKELLSREELTLTAAREQSQAEQLRRREHLAQMELVEFERELTLL 461

RESULT 4

JC2363  
protein kinase (EC 2.7.1.37) ZPK - human  
N;Alternate names: leucine-zipper protein kinase  
C;Species: Homo sapiens (man)  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Feb-2001  
C;Accession: JC2363  
R;Reddy, U.R.; Pleasure, D.  
Biochem. Biophys. Res. Commun. 202, 613-620, 1994  
A;Title: Cloning of a novel putative protein kinase having a leucine zipper domain from  
A;Reference number: JC2363; MUID:94311945; PMID:8037767  
A;Accession: JC2363  
A;Molecule type: mRNA  
A;Residues: 1-668 <RED>  
A;Cross-references: EMBL:U07358  
Experimental source: brain  
Note: the nucleotide sequence for this amino acid sequence is inconsistent with that of  
he codon ACC for residue 661 as Pro, the codon GAACCACTCTCTCCA for residues 664-668 as A  
C;Comment: This protein belongs to the family of non-receptor kinase.  
C;Genetics:  
A;Gene: GDB:ZPK  
A;Cross-references: GDB:383963; OMIM:600447  
A;Map position: 12q13-12q13  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C;Keywords: ATP; leucine zipper; nucleotide binding; P-loop; phosphotransferase  
F;123-371/Domain: protein kinase homology <KIN>  
F;131-139/Region: protein kinase ATP-binding motif  
F;443-471/Region: leucine zipper motif  
F;538-545/Region: nucleotide-binding motif A (P-loop)  
F;152/Active site: Lys #status predicted

Query Match 23.8%; Score 565.5; DB 2; Length 668;  
Best Local Similarity 37.8%; Pred. No. 3.7e-23;  
Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;

QY 10 QIKFDDLQFFENC GGSGSVYRAKWISQDKEVAVKLLKIEKEAEI--LSVLSHRNIIQ 67  
Db 119 EVPFEEILDQWVGSGAQGVFLGRF--HGEEVAVKVRDL-KETDIKHLRKLKHPNIIT 175  
QY 68 FYGVILEPPNYGIVTEYASLGSLYDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPV 127  
Db 176 FKGVCTQAPCYCLIMEFCAQGLYEVLRAGRPVPSL-LVDWSMGIAGMNYLHLH--- 230  
QY 128 KYIHRDLKSRNVVIAADGVLCIDFGASR-FHNHTTHMSLVGTFPMAPEVIQSLPVSET 186  
Db 231 KIHRDLKSPNMLITYDDVVKISDFGTSKELSDKSTKMSFAGTVAMMAPEVIRNEPVSEK 290  
QY 187 CDTYSYGVVLMELTREVPFKGLEQLVAMLVVEKNERLTPSSCPSPFAELHQCWEAD 246  
Db 291 VDIWSFGVLMELLTGEIPIKYDVSSAIIWGVGNSLHLVPSSCPDGFKILLRQCWNSK 350  
QY 247 AKRPSFKQIISILESMSNDT-SLPDKNSFLHNKAEMRCEIATLERLK-----KLE 298  
Db 351 PRNPSFRQILHLHDIASADVLSTPQE--TYFKSQAEWREVKLHFEKIKSEGTCLHRL 408  
QY 299 RDLSPKEQE-----LKERERRLKMEQKLTQSNTPLLPLAARMSEESYFESKTEE 350  
Db 409 EELVMRRREELRHALDIREH----YERKLERANN-LYMELNALMLQ---LELKERE 456

RESULT 5

A55318  
serine/threonine protein kinase (EC 2.7.1.-) DLK - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 24-Sep-1999  
C;Accession: A55318  
R;Holzman, L.B.; Merritt, S.E.; Fan, G.  
J. Biol. Chem. 269, 30808-30817, 1994

A;Title: Identification, molecular cloning, and characterization of dual leucine zipper  
s.  
A;Reference number: A55318; MUID:95074107; PMID:7983011  
A;Accession: A55318  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-888 <HOL>  
A;Cross-references: GB:U14636; NID:g602677; PIDN:AAAS7280.1; PID:g602678  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol  
C;Keywords: ATP; leucine zipper; phosphotransferase  
F;156-404/Domain: protein kinase homology <KIN>  
F;164-172/Region: protein kinase ATP-binding motif

Query Match 23.8%; Score 565.5; DB 2; Length 888;  
Best Local Similarity 37.8%; Pred. No. 5.1e-23;  
Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;

QY 10 QIKFDDLQFFENC GGSGSVYRAKWISQDKEVAVKLLKIEKEAEI--LSVLSHRNIIQ 67  
Db 152 EVPFEEILDQWVGSGAQGVFLGRF--HGEEVAVKVRDL-KETDIKHLRKLKHPNIIT 208  
QY 68 FYGVILEPPNYGIVTEYASLGSLYDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPV 127  
Db 209 FKGVCTQAPCYCLIMEFCAQGLYEVLRAGRPVPSL-LVDWSMGIAGMNYLHLH--- 263  
QY 128 KYIHRDLKSRNVVIAADGVLCIDFGASR-FHNHTTHMSLVGTFPMAPEVIQSLPVSET 186  
Db 264 KIHRDLKSPNMLITYDDVVKISDFGTSKELSDKSTKMSFAGTVAMMAPEVIRNEPVSEK 323  
QY 187 CDTYSYGVVLMELTREVPFKGLEQLVAMLVVEKNERLTPSSCPSPFAELHQCWEAD 246  
Db 324 VDIWSFGVLMELLTGEIPIKYDVSSAIIWGVGNSLHLVPSSCPDGFKILLRQCWNSK 383  
QY 247 AKRPSFKQIISILESMSNDT-SLPDKNSFLHNKAEMRCEIATLERLK-----KLE 298  
Db 384 PRNPSFRQILHLHDIASADVLSTPQE--TYFKSQAEWREVKLHFEKIKSEGTCLHRL 441  
QY 299 RDLSPKEQE-----LKERERRLKMEQKLTQSNTPLLPLAARMSEESYFESKTEE 350  
Db 442 EELVMRRREELRHALDIREH----YERKLERANN-LYMELNALMLQ---LELKERE 489

RESULT 6

T52626  
probable mitogen-activated protein kinase MAP3K delta-1 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 08-Dec-2000  
C;Accession: T52626  
R;Jouanic, S.; Hamal, A.; Lepince, A.S.; Tregear, J.W.; Kreis, M.; Henry, Y.  
Gene 229, 171-81, 1999  
A;Title: Characterisation of novel plant genes encoding MEKK/STE11 and RAF-related prot  
A;Reference number: Z24447; MUID:99196996; PMID:10095117  
A;Accession: T52626  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-406 <JOU>  
A;Cross-references: EMBL:Y14199; NID:g2253009; PIDN:CAA74591.1; PID:g2253010  
A;Experimental source: cultivar Columbia  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

Query Match 23.6%; Score 560.5; DB 2; Length 406;  
Best Local Similarity 41.0%; Pred. No. 4e-23;  
Matches 114; Conservative 55; Mismatches 92; Indels 17; Gaps 6;

QY 10 QIKFDDLQFFENC GGSGSVYRAKWISQDKEVAVKLL-----KIEKAEILSV 59  
Db 129 EIMMEDLQIGERIGISYGEVYRAEW--NGTEVAVKKFLDQDFSGDALTGFKSEIEMLR 186  
QY 60 LSHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSNRSEEMDMHMTWATDVAKGMH 119  
Db 187 LRHPNVVLFMGAVTRPPNFSILTEFLPRGSLYRLH-RPNHQLDEKRRMALDVAKGMN 245  
QY 120 YLHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFHNHT--THMSLVGTFPMAPEV 177



Db 246 YLHTSHPT-VVHRDLKSPNLLVDKNVVKVCDPGLSRMKHHTYLSKSTAGTPEWMADEV 304  
QY 178 IQSLPVSCTDITYSYGVVLMEMLTREVPFKGLEGLQVAMLVKNERLTIPSSCPSPFAE 237  
Db 305 LRNEEPANEKCDVYSFGVILWELATSRVPWKGGLNPMQVGVAFQNRRLIIPDDIDLTVAQ 364  
QY 238 LHOQWEADAKRPSFKQIISILESMSNDTSLPDKCNS 275  
Db 365 IIRECWQTEPHLRPSFTQLMQSLKRLQG-LNISNRANT 401

## RESULT 7

JC5399  
dual leucine zipper kinase (EC 2.7.-.-) - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 07-Jul-1997 #sequence\_revision 18-Jul-1997 #text\_change 28-May-1999  
C;Accession: JC5399  
R;Matsui, N.; Sarkar, G.; Shuto, T.; Marrs, J.; Bronk, J.T.; Mizuno, K.; Bolander, M.E.  
Biochem. Biophys. Res. Commun. 229, 571-576, 1996  
Title: Identification of a dual leucine zipper kinase involved in rat fracture repair.  
Reference number: JC5399; MUID:97127443; PMID:8954939  
A;Accession: JC5399  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-888 <MAT>  
C;Comment: This enzyme is involved in regulating cell function in the musculoskeletal system.  
C;Genetics:  
A;Gene: rDLK  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C;Keywords: ATP; phosphotransferase  
F;54-75, 89-98/Region: glycine-rich  
F;156-454/Domain: kinase catalytic #status predicted <CAT>  
F;156-404/Domain: protein kinase homology <KIN>  
F;164-172/Region: protein kinase ATP-binding motif  
F;421-449/Region: leucine zipper motif  
F;472-500/Region: leucine zipper motif  
F;557-888/Region: glycine-serine-proline rich #status predicted

## Query Match

23.6%; Score 560.5; DB 2; Length 888;  
Best Local Similarity 36.7%; Pred. No. 9.4e-23;  
Matches 133; Conservative 73; Mismatches 131; Indels 25; Gaps 11;

QY 10 QIKPDDLQFFENGCGSGSVYRAKWISQDKEAVAKKLKEKEAEI--LSVLSHRNIQ 67  
Db 152 EYVPEEILDLQWVGSAQGAFLGRF--HGEVAVKKVRDL-KETDIKHLRKLKHPNITT 208  
QY 68 FYGVILPNNYGIIVTEYASLSGLDYINSNRSEMDHMTWATDVAKGMHYLHMEAPV 127  
Db 209 FKGVCTQAPCYCILMEFCAQGLYEVLRAGRPTSL--LVDWSMGIAAGMYIHLH--- 263  
QY 128 KYIHRDLKSRNVVIAADGVLCIDFGASR-FHNHTHMSLVGTFPMAPVIOSLPVSET 186  
Db 264 KIHRDLKSPNMLITYDDVVKISDFGTSKELSDKSTKMSFAGTVAWMAPEVIRNEPVSEK 323  
QY 187 CDITYSGVVLWEMLTREVPFKGLEGLQVAMLVKNERLTIPSSCPSPFAELHQCWEAD 246  
Db 324 VDIWSFGVVLWELLTGEIPKVDSSAIWGVGSNSLHPVPSSCPDGFKILLRQCWNSK 383  
QY 247 AKRPSFKQIISILESMSNDT-SLPDKNSFLHNKAWECEIATLERLKLERLDSFKE 305  
Db 384 PRNRPSPRQIILHLDIASADVLTPEQ--TYFKSQAEWREEVYLFHEKIKSEGTCLHRL 441  
QY 306 QELKERER-----LKWEQKLTQSNTPLLPLAARNSEESYFESK-TEESNSAEM 356  
Db 442 EELVRRRGEGFRHALDIREHYEKLERANN--LYMELNALMLQLLELTERELLRAQALER 499

## RESULT 8

QY 357 SC 358  
Db 500 RC 501

T16747  
hypothetical protein R13F6.7 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Mar-2000

C;Accession: T16747

R;Miller, N.

submitted to the EMBL Data Library, April 1994

A;Description: The sequence of C. elegans cosmid R13F6.

A;Reference number: Z18570

A;Accession: T16747

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-328 <MIL>

A;Cross-references: EMBL:U00046; NID:g470358; PID:g470364; PIDN:AAC47047.1; GSPDB:GN000.

A;Experimental source: strain Bristol N2; clone R13F6

C;Genetics:

A;Gene: CESP:R13F6.7

A;Map position: 3

A;Introns: 20/3; 160/3; 222/2; 286/2

C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3

## Query Match

23.0%; Score 545.5; DB 2; Length 328;  
Best Local Similarity 41.2%; Pred. No. 2e-22;  
Matches 112; Conservative 54; Mismatches 93; Indels 13; Gaps 5;

QY 8 FVQIKPDDLQFFENGCGSGSVYRAKWISQD--KEVAVKLLKIEKEAELSVLSHRN 64  
Db 44 FPDIQRDDIQGDHIGVGTFGAVFSGNWTLPDGSQRTALKVFLKEKEAELSKIRHKN 103  
QY 65 IIOFYGVILEPPN-YGIIVTEYASLSGLDYINSNRSEEM-----DMDHMTWATDVAK 116  
Db 104 IIOFYGICKATGNDFFIVTEYAEKGLDYFIHSEESQSFASSSGGNSFDVYVKMASQIAS 163  
QY 117 GMHYLHMEAPVKVTHRDLSKRNVIADGVLCIDFGASRFHNHT-THMSLVGTFPMAP 175  
Db 164 GIOYLHYDAVDTIHRDLKSKNVLDKNLVCKICDFGTSKDLTHSCTAPSWGTAAMSP 223  
QY 176 E-VIOSLPVSETCDITYSYGVVLMEMLTREVPFKGLEGLQVAMLVKNERLTIPSSCPRS 234  
Db 224 EMILQSEGLTTATDVWSYGVVLMELLSKEVPKYDYSEFRIFTMITQSGITLAIIPSPCAP 283  
QY 235 FAELHQCWEADAKRPSFKQIISILESMSND 266  
Db 284 LKQLMSNCWKMTPRDRANMRQIQGELNRLAGN 315

## RESULT 9

T48544

MAP3K delta-1 protein kinase - Arabidopsis thaliana

N;Alternate names: protein F14F18.20

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C;Accession: T48544

R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft

submitted to the Protein Sequence Database, April 2000

A;Reference number: Z24490

A;Accession: T48544

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-886 <BEV>

A;Cross-references: EMBL:AL163812

A;Experimental source: cultivar Columbia; BAC clone F14F18

C;Genetics:

A;Map position: 5

A;Introns: 148/3; 180/3; 327/3; 362/2; 568/3; 658/1; 671/3; 694/3; 728/2; 754/3; 810/3;

A;Note: F14F18.20

## Query Match

22.3%; Score 529; DB 2; Length 886;  
Best Local Similarity 43.3%; Pred. No. 4.5e-21;  
Matches 107; Conservative 45; Mismatches 79; Indels 16; Gaps 5;

QY 10 QIKPDDLQFFENGCGSGSVYRAKWISQDKEAVAKKL-----KIEKEAELSV 59

Db 643 EIMMEDLQIGERIGISGEVYRAEW--NGTEVAVKKFLDQDFSGDALTOFKSEIEIMLR 700  
QY 60 LSHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEMDMDHIMTWATDYAKGMH 119  
Db 701 LRHPNVLFMGAVTRPPNSILTEFLPRGSLYRLIH-RPNHQLDEKRRMMLDYAKGMN 759  
QY 120 YLHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHT--THMSLVGTFPMNAPEV 177  
Db 760 YLHTSHPT-VVHRDLKSPNLLVDKNWVKVCDPGLSRMKHHTYLSKSTAGTPEWMAPEV 818  
QY 178 IQSLPVSETCDTYSYGVVLMWMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAE 237  
Db 819 LRNEPANCKDVYSFGVILWELATSRVPMWKGILNPMQVGVAVGFQNRRLIIPDIDLTVAAQ 878  
QY 238 LHHQCWE 244  
Db 879 IIRECWO 885

ULT 10  
763  
Hypothetical protein F25P22.8 (imported) - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear crests)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: F96763  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Hutzar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: F96763  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-1030 <STO>  
A/Cross-references: GB:AE005173; NID:g6692730; PIDN:AAF24836.1; GSPDB:GN00141  
C/Genetics:  
A/Gene: F25P22.8  
A/Map position: 1

Query Match 21.5%; Score 510.5; DB 2; Length 1030;  
Best Local Similarity 38.0%; Pred. No. 5.2e-20;  
Matches 109; Conservative 54; Mismatches 93; Indels 31; Gaps 7;  
QY 10 QIKFDDLOFFENCGGSGFSGVYRAKWISODKEVAVKKL-----KIEKEAELISV 59  
Db 742 EILWEEITVGERIGLSYGEVYRGDW--HGTEVAVKKFLDQDLTGEALEEFRSEVRIMKK 799  
QY 60 LSHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEMDMDHIMTWATDYAKGMH 119  
Db 800 LRHPNVLFMGAVTRPPNSILTEFLPRGSLYRLIH-RPNHQLDERRLMALDARGMN 858  
QY 120 YLHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTHM--SLVGTFFPMNAPE 176  
Db 859 YLHSCNPM-IVHRDLKSPNLLVDKNWVKVCDPGLSRM-KHSTYLSKSTAGTAENMAPE 916  
QY 177 VIQSLPVSETCDTYSYGVVLMWMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFA 236  
Db 917 VLNEPADKCDVYSYGVILWELFTLQOPWGMKNPMQVGVAVGFQHRRLDIPDFVDPALIA 976  
QY 237 ELLHQCWEADAKRPSFKQIISILESMS-----NDTSLP 270  
Db 977 DLISKWQTDSTKLSPFAEIMASLKRLOKPTGSGNIPRPVSSSSSLP 1023

RESULT 11  
B87950  
protein F33E2.2 (imported) - Caenorhabditis elegans

C/Species: Caenorhabditis elegans  
C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C/Accession: B87950  
R/Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo  
A/Reference number: A75000; MUID:99069613; PMID:9851916  
A/Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_el  
A/Note: published extra appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an  
A/Accession: B87950  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-848 <STO>  
A/Cross-references: GB:chr\_I; PIDN:CAB06544.1; PID:g3876653; GSPDB:GN00019; CESP:F33E2.  
C/Genetics:  
A/Gene: F33E2.2  
A/Map position: 1

Query Match 21.3%; Score 505; DB 2; Length 848;  
Best Local Similarity 30.4%; Pred. No. 8.3e-20;  
Matches 145; Conservative 69; Mismatches 159; Indels 104; Gaps 14;  
QY 10 QIKFDDLOFFENCGGSGFSGVYRAKWISODKEVAVKKLIEKEAEI--LSVLSHRNIIQ 67  
Db 56 EIPFDAISELWELSGSGGAVFRGQ--LENTVAVKVKNQL-KETELKHLRLHQNIIIE 112  
QY 68 FYGVILEPPNYGIVTEYASLSGLDYINSNRSEMDMDHIMTWATDYAKGMHYLHMEAPV 127  
Db 113 FLGVCSKSPCYCIWMEYCSKGQLCTVLKSRNTITREL--FAQWKEIADGMYLHQN--- 167  
QY 128 KYIHRDLKSRNVVIAADGVLCICDFGASRFHN--HTHMSLVGTFPMNAPEVLSLVSE 185  
Db 168 KYIHRDLKSPNLLISAEDSIKICDFGTSIMQKMDSTMSPCGTSMWMAPEMLKKQPCNE 227  
QY 186 TCDTYSYGVVLMWMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELHQCWEA 245  
Db 228 KVDVYSFGVILWMLTRETPYANIAQMAIIFGVGTNILLSMPBEAPKGLVLLIKQCLSQ 287  
QY 246 DAKRPSFKQI-----ISILESMSNDTSLPD-- 271  
Db 288 KGRNRPSFSHIRQWEIIFKPELFEMTEBEWQLAWDYREFAKCIOYPSYVTRDHGCPKSA 347  
QY 272 -----KCNSEFLHNKAEMRCEIATLERLKLERDLSFKEQLKERERRLKMEQKL 322  
Db 348 FAMEEIIQRKRHEQLNHIKDIRMYEMKLKRTNKNYDKLQGCFTELKKESELAWEKDL 407  
QY 323 TEQ-----SNTP--LLPLAARMS--EESYFESKTEES---NSAEMSCQITATSNGECH 369  
Db 408 TEREQWNNQNSPKAVAPRAQLRGYPNEGIDMSDEDVQPCRGSPYRCSNTSSS----- 462  
QY 370 GNNPSLQAMMLMGFDIFSMNKAGAVMHSQMIAKONSSKITS---KRGKKYN 423  
Db 463 -----SGVQSSPFSRQSSSSSSAGQOTRREGAN 491

RESULT 12  
T20082  
Hypothetical protein F33E2.2 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C/Accession: T20082; T21703  
R/Lennard, N.  
submitted to the EMBL Data Library, April 1998  
A/Reference number: Z19221  
A/Accession: T20082  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-855 <WIL>  
A/Cross-references: EMBL:AL022593; PIDN:CAA18635.2; GSPDB:GN00019; CESP:F33E2.2  
A/Experimental source: clone C49G9  
R/Lennard, N.  
submitted to the EMBL Data Library, January 1997  
A/Reference number: Z19461



```

OY      236 AETLHQCEWADAKRPSFKQIISILESM 263
      : : : : : : : : : : : : : : : :
Db      781 AAIEGCTNEPWRKRPSPFATIMDLRPL 808

```

## RESULT 15

JC5955  
 transforming growth factor-beta activated kinase (EC 2.7.--.) 1a - human  
 C;Species: Homo sapiens (man)  
 C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
 C;Accession: JC5955  
 R;Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.  
 Biochem. Biophys. Res. Commun. 243, 545-549, 1998  
 A;Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind  
 A;Reference number: JC5955; MUID:98153801; PMID:9480845  
 A;Accession: JC5955  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-579 <SAK>  
 A;Cross-references: DDBJ:AB009356; NID:g2924623; PIDN:BA25025.1; PID:g2924624  
 A;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
 C;Keywords: phosphotransferase

|                           |       |                    |            |             |
|---------------------------|-------|--------------------|------------|-------------|
| Query Match               | 20.3% | Score 483;         | DB 2;      | Length 579; |
| Best Local Similarity     | 30.2% | Pred. No. 8.1e-19; |            |             |
| Matches 131; Conservative | 82;   | Mismatches 151;    | Indels 70; | Gaps 18;    |

```

QY 7 SFWQIKEDDLQFENNCGGSGFSGSVYRAKNI SODKEVAVKKLLKIEKEAE-----IL 57
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 27 NFEEIDYKEIEVEEVGRGAFGVCKAKW--RAKDVAIK--QIESESEKAFIYELRQL 81
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 58 SVLSHRNIIQFYGVILEPPNNGIVTEYASLGSLYDIINSNRSEEM--DMDHMTWATDV 114
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 82 SRVNHPNIVKLYGACLNP--VCLVMEYAEGGSLYNVLHG--AEPLPYTTAAHAMSWCLOQ 137
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 115 AKGMHYLHMEAPVKVIHRDLKSRNVIAADG-VLKICDFG-ASRFHNHTHMSLVGTFFPW 172
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 138 SOGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFGTACDIQTHMTNNK--GSAAW 195
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 MABEVIQSLPVSETCDTYSYGVVLWEMLTREVPEFKGLEG--LOVAMLVEKNERLTIPSS 230
    ||||| : ||| : : : : : : : : : : : : : : : : : : : : : :
Db 196 MABEVEFGSNYSEKCDVFSWGIILWEVITRKPFEIDIGPAFRIMW-AVHNGTRPPLIKN 254
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 231 CPRSFAELLHQCWEADAKRPSFKQIISIL-----ESMSNDT- 267
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 255 LKPRIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPGADEPLQYPCQYSDGQSN SAT 314
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
268 ---SLPDKCNSEFLHNKAEMWRC-IEATLERLKKLERDLSFKEQELKERERRLKNWEQKLT 323
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 315 STGSRMDIASTNTSNKSDTNMEQVPATNDTIKRLESKLLKNQAKQOSESGRLSLGASRG 374
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 324 EQSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQITATSNNGEGHGMNPSLQAMLMGF 383
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 375 SVEBUP---PTS-----EGKRMADMSIEIARIAATT-GNGQPRRRSIODLTVTGT 421
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 384 --GDIFSMNKAGAV 395
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 422 EPGOVSSRRSSSPSV 435
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: May 1, 2003, 20:38:37  
Job time : 37.2873 secs





GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:13:36 ; Search time 22.8452 Seconds

(without alignments)  
826.070 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 2375  
Sequence: 1 MSSLGASFVQIKFDDLQFFE.....GDDDDDDGEEEDNDMNSE 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID           | Description        |
|------------|-------|-------------|--------|--------------|--------------------|
| 1          | 601   | 25.3        | 954    | 1 M3KA_HUMAN | Q02779 homo sapien |
| 2          | 593   | 25.0        | 394    | 1 M3X9_HUMAN | P80192 homo sapien |
| 3          | 565.5 | 23.8        | 859    | 1 M3KC_HUMAN | Q12852 homo sapien |
| 4          | 565.5 | 23.8        | 888    | 1 M3KC_MOUSE | Q60700 mus musculu |
| 5          | 563.5 | 23.7        | 888    | 1 M3KC_RAT   | Q63796 rattus norv |
| 6          | 488   | 20.5        | 821    | 1 CTRL_ARATH | Q05609 arabidopsis |
| 7          | 483   | 20.3        | 579    | 1 M3K7_MOUSE | Q62073 mus musculu |
| 8          | 471.5 | 19.9        | 606    | 1 M3K7_HUMAN | Q43318 homo sapien |
| 9          | 434   | 18.3        | 746    | 1 ABL_MLVAB  | P00521 abelson mur |
| 10         | 410   | 17.3        | 1130   | 1 ABL1_HUMAN | P00519 homo sapien |
| 11         | 405   | 17.1        | 1123   | 1 ABL1_MOUSE | P00520 mus musculu |
| 12         | 405   | 17.1        | 1182   | 1 ABL2_HUMAN | P42684 homo sapien |
| 13         | 402   | 16.9        | 1520   | 1 ABL_DROME  | P00522 drosophila  |
| 14         | 401.5 | 16.9        | 1584   | 1 KYK1_DICDI | P18160 dictyosteli |
| 15         | 385.5 | 16.2        | 439    | 1 ABL_FSVHY  | P10447 feline sarc |
| 16         | 385.5 | 16.2        | 536    | 1 FYN_XIPHE  | P27446 xiphophorus |
| 17         | 379.5 | 16.0        | 505    | 1 FRK_HUMAN  | P42685 homo sapien |
| 18         | 378.5 | 15.9        | 393    | 1 M3K7_DROME | P83104 drosophila  |
| 19         | 378.5 | 15.9        | 536    | 1 FYN_XENLA  | P13406 xenopus lae |
| 20         | 376.5 | 15.9        | 536    | 1 FYN_HUMAN  | P06241 homo sapien |
| 21         | 374   | 15.7        | 410    | 1 KYK2_DICDI | P18161 dictyosteli |
| 22         | 369.5 | 15.6        | 1196   | 1 ABL1_CAEBL | P03949 caenorhabdi |
| 23         | 365.5 | 15.4        | 535    | 1 YRK_CHICK  | Q02977 gallus gall |
| 24         | 359.5 | 15.1        | 533    | 1 FYN_MOUSE  | P39688 mus musculu |
| 25         | 358.5 | 15.1        | 587    | 1 SRC_AVIS2  | P15054 avian sarco |
| 26         | 358.5 | 15.1        | 675    | 1 BMX_HUMAN  | P51813 homo sapien |
| 27         | 358   | 15.1        | 822    | 1 FGRI_HUMAN | P13362 homo sapien |
| 28         | 358   | 15.1        | 822    | 1 FGRI_MOUSE | P16092 mus musculu |
| 29         | 357   | 15.0        | 552    | 1 SRC1_DROME | P00528 drosophila  |
| 30         | 357   | 15.0        | 819    | 1 FGRI_CHICK | P21804 gallus gall |
| 31         | 355.5 | 15.0        | 517    | 1 FGR_MOUSE  | P14234 mus musculu |
| 32         | 355.5 | 15.0        | 531    | 1 SRC1_XENLA | P13115 xenopus lae |
| 33         | 355.5 | 15.0        | 531    | 1 SRC2_XENLA | P13116 xenopus lae |

|    |       |      |      |              |                    |
|----|-------|------|------|--------------|--------------------|
| 34 | 354   | 14.9 | 497  | 1 SPK1_DUGTI | P42687 dugesia tig |
| 35 | 353.5 | 14.9 | 532  | 1 SRC_CHICK  | P00523 gallus gall |
| 36 | 353   | 14.9 | 822  | 1 FGRI_RAT   | Q04589 rattus norv |
| 37 | 352   | 14.8 | 806  | 1 CEK2_CHICK | P18460 gallus gall |
| 38 | 351.5 | 14.8 | 557  | 1 SRC_AVIST  | P14085 avian sarco |
| 39 | 351.5 | 14.8 | 757  | 1 HT16_HYDAT | P53356 hydra atten |
| 40 | 350.5 | 14.8 | 527  | 1 TXK_HUMAN  | P42681 homo sapien |
| 41 | 350.5 | 14.8 | 533  | 1 FYN_CHICK  | Q05876 gallus gall |
| 42 | 350   | 14.7 | 806  | 1 FGRI_HUMAN | P22607 homo sapien |
| 43 | 350   | 14.7 | 1426 | 1 EGFR_DROME | P04412 drosophila  |
| 44 | 349.5 | 14.7 | 535  | 1 SRC_HUMAN  | P12931 homo sapien |
| 45 | 349.5 | 14.7 | 568  | 1 SRC_AVISS  | P14084 avian sarco |

## ALIGNMENTS

```
RESULT 1
M3KA_HUMAN
ID M3KA_HUMAN STANDARD; PRT; 954 AA.
AC Q02779; Q12761, Q14871;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 10 (EC 2.7.1.37)
DE (Mixed lineage kinase 2) (Protein kinase MST).
GN MAP3K10 OR MLK2 OR MST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96128179; PubMed=8536694;
RA Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,
RA Sutherland G.R., Simpson R.J.;
RT "Complete nucleotide sequence, expression, and chromosomal
RT localisation of human mixed-lineage kinase 2.";
RL Eur. J. Biochem. 234:492-500 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95249256; PubMed=7731697;
RA Katoh M., Hirai M., Sugimura T., Terada M.;
RT "Cloning and characterization of MST, a novel (putative)
RT serine/threonine kinase with SH3 domain.";
RL Oncogene 10:1447-1451 (1995).
RN [3]
RP SEQUENCE OF 244-480 FROM N.A.
RC TISSUE=Colon epithelium;
RX MEDLINE=93238756; PubMed=8477742;
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;
RT "Identification of a new family of human epithelial protein kinases
RT containing two leucine/isoleucine-zipper domains.";
RL Eur. J. Biochem. 213:701-710 (1993).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X90846; CAA62351.1; -
DR EMBL; Z48615; CAA88531.1; -
DR PIR; S32468; S32468.
```

DR HGSP; P11362; 1FGK.  
DR Genew; HGNC:6849; MAP3K10.  
DR MIM; 600137; -  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00018; SH3; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00221; STYKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS50002; SH3; 1.  
DR Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase; ATP-binding; SH3 domain.  
KW ATP-binding; SH3 domain.  
FT DOMAIN 2 5  
FT DOMAIN 16 81  
FT DOMAIN 98 360  
FT NP\_BIND 104 112  
FT BINDING 125 125  
FT ACT\_SITE 222 222  
FT DOMAIN 384 405  
FT DOMAIN 419 440  
FT DOMAIN 449 463  
FT DOMAIN 462 464  
FT CONFLICT 465 480  
FT CONFLICT 471 471  
FT CONFLICT 807 807  
FT CONFLICT 818 818  
SQ SEQUENCE 954 AA; 103623 MW; 538F4AAA559B0ABA CRC64;  
Query Match 25.3%; Score 601; DB 1; Length 954;  
Best Local Similarity 35.4%; Pred. No. 7.5e-32;  
Matches 155; Conservative 63; Mismatches 160; Indels 60; Gaps 10;  
QY 10 QIKEDDLQFFENC GGSGFSGVYRAKVISQDKEVAVKKL-LKIEK-----EAEIL 57  
DB 92 EIPFHELQLEIITGVGFGKVRALW--RGEVAVAKARLDEKDPAVTAQVCQEARLF 149  
QY 58 SVLSHRNIIQFYGVILEPPNYGIYTEYASLSGLYDYINSNRSEMDMDHIM-TWATDVAK 116  
DB 150 GALQHPNIIALRGACLNPHLCVMEYARAGGALSRLVAGRVP----PHVLVNWAVQVAR 205  
QY 117 GMYILMEAPVKVIHRDLKSRNVI-----AADGVLCIDFGASRFHNHTHMSLVG 168  
DB 206 GMYILHNDAPVPIIHRDLKSNILILEAIEHNHLADTVLKITDFGLAREMHKTKMSAAG 265  
QY 169 TFPWMADEVIOQLPVSETCDITYSYGVVLWEMLTREVPFKGLGLOVAVWLVEKNERLTIP 228  
DB 266 TYAWMADEVIRLSLFSKSSDVMSFVLLWELLTGEVPRYREIDALAVAYGVAMNKLTLPIIP 325  
QY 229 SSCPRSPAEILLQCWEADAKKRPSEKQIISILESMSNDTSLPDKNSFLHNKAWRCEIE 288  
DB 326 STCEPPEFARLLEECWDPDPHGRPDGSIKRLLEVEIQSALFQMPLESHSLQEDWKLEIQ 385  
QY 289 ATLERLKLIERDLSPKEQL-----KERERLKMWEQKLTEQSNTPLLPLAARMSEE 341  
DB 386 HMFDDLRTKEKELRSREELLRAAQEQRFQEQLRREQELAEKEMDIVEREHLHLMCQL 445  
QY 342 SYFESKTEESNSAEMSCQITATSNGBGHGMNPSLQAMMLMGFGDIFSMNKAGAVMHSGMQ 401  
DB 446 SQEKPRVRRKRGKGNFKRSRLKLREGGSHISLPS-----GFEH-----K 483  
QY 402 INMQAKONSSKTSKRRG 419

DB 484 ITVQA----SPTLDKRGK 497  
RESULT 2  
M3K9\_HUMAN STANDARD; PRT; 394 AA.  
AC P80192;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Mitogen-activated protein kinase kinase kinase 9 (EC 2.7.1.1-) (Mixed lineage kinase 1) (Fragment).  
DE MAP3K9 OR MLK1 OR PRK1.  
GN Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCB1\_TaxID=9606;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Colon epithelium;  
RC MEDLINE=93238756; PubMed=8477742;  
RX Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;  
RA "Identification of a new family of human epithelial protein kinases containing two leucine/isoleucine-zipper domains."  
RT Eur. J. Biochem. 213:701-710(1993).  
RL -1- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF COLONIC, BREAST AND OESOPHAGEAL ORIGIN.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC MAP KINASE KINASE KINASE SUBFAMILY.  
CC PIR; S32467; S32467.  
DR PIR; JU0229; JU0229.  
DR HSSP; P12931; 1FMK.  
DR Genew; HGNC:6861; MAP3K9.  
DR MIM; 600136; -  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00219; TyKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase; ATP-binding.  
KW NON\_TER 1 1  
FT DOMAIN 3 271  
FT NP\_BIND 9 17  
FT BINDING 30 30  
FT ACT\_SITE 127 127  
FT DOMAIN 289 310  
FT DOMAIN 324 345  
FT DOMAIN 354 368  
SQ SEQUENCE 394 AA; 44975 MW; DBE40B7D31047FD8 CRC64;  
Query Match 25.0%; Score 593; DB 1; Length 394;  
Best Local Similarity 39.3%; Pred. No. 8.2e-32;  
Matches 133; Conservative 59; Mismatches 114; Indels 32; Gaps 5;  
QY 15 DLQFFENC GGSGFSGVYRAKVISQDKEVAVK-----KLKIEKEAEILSVLSH 62  
DB 2 ELTLEEIIGIGFGKVRFAWIGD--EVAVKARHDPDEDISQTIENVRQEKLFAMLKH 59  
QY 63 RNIIQFYGVILEPPNYGIYTEYASLSGLYDYINSNRSEMDMDHIMTWATDVAKGMYILH 122  
DB 60 PNIALRGVCLKEPNLCVMEFARGGPLNRVLSGR--IPDILVNWAVQIARGWNYLH 116  
QY 123 MEAPVKVIHRDLKSRNVI-----AADGVLCIDFGASRFHNHTHMSLVGTFPWWA 174  
DB 117 DEAIVPPIIHRDLKSSNIILOQVENGDLSNKLKITDFGLAREWHRTTKMSAAGTYAWMA 176  
QY 175 PEVIQSLPVSETCDITYSYGVVLWEMLTREVPFKGLGLOVAVWLVEKNERLTIPSSCPRS 234

Db 177 PEVIRASMSKSDVWSYGVLLMELLTGEVPRFGIDGLRVAVGVAMNKLALPIPTCPPEP 236  
 QY 235 FAELHOCWEADAKRPSFKQIISLESMSNDTSLPDKCNSEFLHNKAEMRCEIATLERL 294  
 Db 237 FAKMEDCWNPDPHSRPSFTNILDQLTTEESGFPEMPKDSFHCLODNWKHEIQEMFDQL 296  
 QY 295 KLEERDLSFKEOE-----KERERRLKMEQKLTQEQ 325  
 Db 297 RAKEKELRTWEELTRALQOKNOEELLRRREQLAER 334

RESULT 3  
 M3KC\_HUMAN STANDARD; PRT; 859 AA..  
 ID M3KC\_HUMAN  
 AC Q12852;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)  
 DE (Leucine-zipper protein kinase) (ZPK).  
 GN MAP3K12 OR ZPK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Teratocarcinoma;  
 RX MEDLINE=94311945; PubMed=8037767;  
 RA Reddy U.R.; Pleasure D.;  
 RT "Cloning of a novel putative protein kinase having a leucine zipper  
 domain from human brain."  
 RL Biochem. Biophys. Res. Commun. 202:613-620(1994).  
 CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.  
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in  
 CC vitro.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- COFACTOR: Magnesium.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By  
 CC similarity).  
 CC -1- TISSUE SPECIFICITY: Highly expressed in brain and kidney.  
 CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol  
 CC under basal conditions and dephosphorylated when membrane-  
 CC associated (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE KINASE SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U07358; AAA67343.1; -  
 DR HSSP; P12931; 1FMK.  
 DR Genew; HGNC:6851; MAP3K12.  
 DR MIM; 600447; -  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
 KW phosphorylation; Magnesium; Membrane.  
 FT DOMAIN 125 366 PROTEIN KINASE.  
 FT NP\_BIND 131 139 ATP (BY SIMILARITY).  
 FT BINDING 152 152 ATP (BY SIMILARITY).

FT ACT\_SITE 236 236 BY SIMILARITY.  
 FT DOMAIN 665 668 POLY-PRO.  
 FT DOMAIN 720 725 POLY-GLU.  
 SQ SEQUENCE 859 AA; 93188 MW; 0E5209792C5C6F05 CRC64;

Query Match 23.8%; Score 565.5; DB 1; Length 859;  
 Best Local Similarity 37.8%; Pred. No. 1.4e-29;  
 Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;

QY 10 QIKFDLQFFENCGGSGFSVYRAKWSODKEVAVKLLKIEKAEI--LSVSHRNITQ 67  
 Db 119 EVPFEEILDQWVGSGAGAVFLGRF--HGEVAVKVRDL-KETDIKHLRLKHPNIIT 175  
 QY 68 FYGVILEPPNYGIVTEYASLSGLDYINSRSEMDMDHMTWATDVAKGMYLHMEAPV 127  
 Db 176 FRGVCTQAPCYCIIMEFCAQGLYELVRAGRPVTPSL--LVDWSMGJAGGMYLHLH--- 230  
 QY 128 KVIHRDLKSRNVIAADGVLCIDFGASR-FHNHTTMSLVGTFPMWAPVYISLPVSET 186  
 Db 231 KIIHRDLKSPNMLITYDDVKISDFGTSKESDKSTKMSFAGTVAMWAPVIRNEPVSEK 290  
 QY 187 CDTYSYGVLLWMLTREVPFEGLEGQVAMVVEKNERLTIPSSCPRSFAELLHOCWEAD 246  
 Db 291 VDIWSFGVLLWELLTGEIRPYKDVDSAIIMWVGNSLHLVPVSSGCPDGFKILLRQCWNSK 350  
 QY 247 AKRPSFKQIISLESMSNDT-SLPDKCNSEFLHNKAEMRCEIATLERLK-----KLE 298  
 Db 351 PNRPSFRQILHLHDIASADVLSTPQE--TYFKSAEMREVEVKLHEKIKSEGTCLHRL 408  
 QY 299 RDLSEFKEOE-----LKERERRLKMEQKLTQEQNTPLPLPLAARMSEESYFESKTEE 350  
 Db 409 EELVMRRREELRHALLDIREH----YERKLERANN--LYMELNALMLQ--LELKERE 456

RESULT 4  
 M3KC\_MOUSE STANDARD; PRT; 888 AA..  
 ID M3KC\_MOUSE  
 AC Q60700; P70286;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)  
 DE (Leucine-zipper protein kinase) (ZPK) (Dual leucine zipper bearing  
 DE kinase) (DLK).  
 GN MAP3K12 OR ZPK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CD-1; TISSUE=Brain;  
 RX MEDLINE=95074107; PubMed=7983011;  
 RA Holzman L.B.; Merritt S.E.; Fan G.;  
 RT "Identification, molecular cloning, and characterization of dual  
 RT leucine zipper bearing kinase. A novel serine/threonine protein kinase  
 RT that defines a second subfamily of mixed lineage kinases."  
 RL J. Biol. Chem. 269:30808-30817(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ICR X Swiss Webster; TISSUE=Brain;  
 RX MEDLINE=96365388; PubMed=8769565;  
 RA Blouin R.; Beaudoin J.; Bergeron P.; Nadeau A.; Grondin G.;  
 RT "Cell-specific expression of the ZPK gene in adult mouse tissues."  
 RL DNA Cell Biol. 15:631-642(1996).  
 RN [3]  
 RP PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.  
 RX MEDLINE=96279269; PubMed=8663324;  
 RA Mata M.; Merritt S.E.; Fan G.; Yu G.G.; Holzman L.B.;  
 RT "Characterization of dual leucine zipper-bearing kinase, a mixed  
 RT lineage kinase present in synaptic terminals whose phosphorylation  
 RT state is regulated by membrane depolarization via calcineurin."  
 RL J. Biol. Chem. 271:16888-16896(1996).





FT DOMAIN 668 671 POLY-PRO.  
FT DOMAIN 698 701 POLY-PRO.  
FT DOMAIN 753 758 POLY-GLU.  
SQ SEQUENCE 888 AA; 96307 MW; 52AD964006BAE149 CRC64;

Query Match 23.7%; Score 563.5; DB 1; Length 888;  
Best Local Similarity 37.8%; Pred. No. 1.9e-29;  
Matches 135; Conservative 70; Mismatches 117; Indels 35; Gaps 13;

QY 10 QIKFDDIQFFENGCGGSGFSVYRAKWISQDKEVAVKLLKIEKEAEI--LSVLSHRNIIQ 67  
DB 152 EVPFEEILDQWVGSGAOGAVFLGRF--HGEVAVKVRDL-KETDIKHLKRLKHPNIIIT 208  
QY 68 FYGVILEPPNYGIVTEYASLSGLYDYNINRSEEMDMHMTATVAKGMHYLHMEAPV 127  
DB 209 FKGVCTQAPCCILMEFCAOGQLYEVLBAGRPTPSL--LVDSMGIAAGMNYLHLH--- 263  
QY 128 KVIHRDLKSRNVVIAADGVLCIDFGASR-FHNHTHMSLVGTFPMWAPVIOQLPVSET 186  
DB 264 KIIHRDLKSRNMLITYDDVVKISDFGTSKELSDKSTKMSFAGIVAMWAPVIRNEPVSEK 323  
QY 187 CDYSGYVLLWMLTREVFPFKGLEGOVAMLVVEKNERLTTPSCPRSPFAELHQCWEAD 246  
DB 324 VDIWSFGVLLWELLTGEIPIYKDVDSAIWGVGNSLHLPVSSCPDGFKILLRQCWNRX 383  
QY 247 AKRPSFKQIISLESMSNDT-SLPKXNSFLHNKAERCEIATLERLK-----KLE 298  
DB 384 PRNRPSPROILLHLDIASADVLSTPQE--TYFKSQAEWREEVKLFHEKIKSEGTCLHRL 441  
QY 299 RDLSPKQOE-----LKERERLKMWEQKLTQSNTPLLPLAARMSEESYFESKTEE 350  
DB 442 EELVWRREELRHALDIREH-----YERKLERANN--LYMELNALMLQ---LELKERE 489

RESULT 6  
CTRL1 ARATH STANDARD; PRT; 821 AA.  
ID CTRL1 ARATH  
AC O05609;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Serine/threonine-protein kinase CTRL1 (EC 2.7.1.37).  
GN CTRL1 OR AT5G03730 OR F17C15\_150.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
NCBI\_TaxID=3702;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=cv. Columbia; TISSUE=Seedling;  
RC MEDLINE=93161417; PubMed=8431946;  
RX MEDLINE=93161417; PubMed=8431946;  
RA Kieber J.J., Rothenberg M., Roman G., Feldmann K.A., Ecker J.R.;  
RT "CTRL1, a negative regulator of the ethylene response pathway in  
RL Arabidopsis, encodes a member of the raf family of protein kinases."  
RN Cell 72:427-441(1993).  
[2]  
SEQUENCE FROM N.A.  
STRAIN=cv. Columbia;  
RC MEDLINE=21016721; PubMed=11130714;  
RX Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,  
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,  
RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,  
RA Nakazaki N., Naito K., Okumura S., Shinpo S., Takeuchi C., Wada T.,  
RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,  
RA Huang B., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,  
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,  
RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,  
RA Belter E., Cordun H., Cordes M., Courtney L., Courtney W., Dante M.,  
RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,  
RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Stromatt C.,  
RA Wagner-McPherson C., Woliam A., Yokum M., Bell M., Dedhia N.,  
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,

RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,  
RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,  
RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,  
RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,  
RA Langham S.-A., McCullagh B., Robben J., Grymoprez B., Zimmermann W.,  
RA Ramsberger U., Wedler E., Balke K., Wedler E., Peters S.,  
RA van Staveren M., Dirke W., Mooijman P., Klein lankhorst R.,  
RA Weltzenegger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,  
RA Bents O., Lemcke K., Kolesov G., Villarroel R., Gielen J., Ardiiles W.,  
RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;  
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis  
thaliana.";  
RL Nature 408:823-826(2000).  
CC -1- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE  
PATHWAY.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.  
CC -1- MISCELLANEOUS: CTRL1 MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,  
RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED  
COYLEDON GROWTH IS IMPAIRED.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
MIL/RAF SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; L08789; AAA32779.1; -  
DR EMBL; L08790; AAA32780.1; -  
DR EMBL; AL162506; CAB82938.1; -  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00221; STYK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding.  
FT DOMAIN 65 69  
FT DOMAIN 135 141  
FT DOMAIN 551 809  
FT NP\_BIND 557 565  
FT BINDING 578 578  
FT ACT\_SITE 676 676  
FT MUTAGEN 596 596  
FT MUTAGEN 694 694  
FT MUTAGEN 694 694  
FT SEQUENCE 821 AA; 90306 MW; 2922D3DCD0CC15BC CRC64;

Query Match 20.5%; Score 488; DB 1; Length 821;  
Best Local Similarity 40.7%; Pred. No. 1.5e-24;  
Matches 109; Conservative 44; Mismatches 99; Indels 16; Gaps 5;  
QY 9 VQIKFDDIQFFENGCGGSGFSVYRAKWISQDKEVAVKLLKIE-----KEAEIIS 58  
DB 544 MDIPWCDLNKIKGAGSGFYHRAEWHGSD--VAVKILMEDQFHAERVNEFLREVAIMK 601  
QY 59 VLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYN--SNRSEEMDMHMTATVAKG 117  
DB 602 RLHRPNIVLFMGAVTQPNLSITVEYLSRGLYRLHLKSGAREQJDERRRRLSMAYDAK 661  
QY 118 MHYLHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRHHNT--THMSLVGTFPMWAP 175  
DB 662 MNYLHNRNP-PIVHRDLKSPNLLVDKYYTVKVCDFGLSLRLKASTFLSSKSAAGTPEWMA 720

```
QY      176 EVIQLPVSETCDTVSYGAVLWMLTREVPFKLEGIOVAMLVKNERLTIPSSCPRSF   235  
       ||::|::||::||::||::||::||::||::||::||::||  
DB     721 EVL RDEPSNEKSVDVSFGIVIMELATLQQPWNGLNPAQVAAGVKCKRLTEPRNLNPQV    780  
  
QY      236 AELLHQCWEADAKKRPFKFKOISILESM 263  
       |::||::|||||::|:  
DB     781 AAIEGCWTNEPWKRFPSPFATIMDLLRPL 808  
  
RESULT 7  
M3K7_MOUSE STANDARD; PRT; 579 AA.  
ID_M3K7_MOUSE Q62073;  
DT_16-OCT-2001 (Rel. 40, Created)  
DT_16-OCT-2001 (Rel. 40, Last sequence update)  
DT_16-OCT-2001 (Rel. 40, Last annotation update)  
DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)  
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-  
DE activated kinase 1).  
GN MAP3K7 OR TAK1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=96123277; PubMed=8533096;  
RX Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,  
RA Taniguchi T., Nishida E., Matsumoto K.;  
RT "Identification of a member of the MAPKK family as a potential  
mediator of TGF-beta signal transduction."  
RL Science 270:2008-2011(1995).  
CC -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.  
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B  
CC ACTIVATION.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC MAP KINASE KINASE SUBFAMILY.  
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outpost at  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
DR EMBL; D76446; BAAl1184.1; -.  
DR HSSP; P08631; LADS.  
DR GMD; MG1:1346877; Map3k7.  
DR InterPro; IPRO00719; Euk_pkinase.  
DR InterPro; IPRO04040; STY_pkinase.  
DR InterPro; IPRO02290; Ser_thr_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk_pkinase; 1.  
DR SMART; SMO0221; STYKC; 1.  
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.  
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.  
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding.  
FT DOMAIN 8 16 POLY-SER.  
FT NP_BIND 36 291 PROTEIN_KINASE.  
FT BINDING 42 50 ATP (BY SIMILARITY).  
FT BINDING 63 63 ATP (BY SIMILARITY).  
FT ACT_SITE 156 156 BY SIMILARITY.  
SQ SEQUENCE 579 AA; 64227 MW; 97C8FEF3C8E283EE CRC64;
```

Query Match

Best Local Similarity    20.3%; Score 483; DB 1; Length 579;

Matches     131; Conservative       82; Mismatches     151; Indels              70; Gaps                  18;

```

Db 27 NFEIDYKEIEVEEVVGRGAFGVCKAKW--RAKDAVK--QIESESEKKAFFIVELRQL 81
QY 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLGSLDYINSNRSEEM---DMDHMTWATDV 114
Db 82 SRVNHNPVTKLYGACLN--VCLVMEYZAEGSLYNVLHG--AEPLPYTAHAAMSWCLOQ 137
QY 115 AKGMHYLHMEAPVYVYIHRDLKSRNVVIAADG-VLKICDFG-ASRFNHTTHMSLVGTFPW 172
Db 138 SQGVAYLHSMQPKALIHRLDKPNNLLVAGTYVLCIDFGTACDIQTHMTNNK--GSAAM 195
QY 173 MAPEVIQSLPVSSETCDTYSYGVVLWEMLTREYVPEKLEG--LOVAMLVYEKNERLTISS 230
Db 196 MAPEVFEKSNYSEKCDVFSWGIITLMEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
QY 231 CPRSFAELLHQCEWADAKKRPSFKQIISIL-----ESMSNDT- 267
Db 255 LPKPIESLMTRCWSKDPSPQSPSMEEIVKIMTHLMRYFPGADEPLQYPCQYSDGQNSAT 314
QY 268 ---SLPDKCNFLHNKAWRCE-IEATLERLKKLERDLSFKEQLKERERRLKWEQKLT 323
Db 315 STGSFMDIASTNTSNKSDPTNMEQVPATNDTIKRLSKLLKNQAKQOSESGRLSLGASRGS 374
QY 324 EOSNTPLLPLAARMSESYFESKTEESNSAEMSCQITATNGEGHGMNPSLOAMLMGF 383
Db 375 SVESLP--PTS-----EGKMSADMSIEARIVATA-GNGQPRRSIQLDVTGT 421
QY 384 --GDIFSMNKAGAV 395
Db 422 EPGQVSSRRSSPSV 435

RESULT 8
M3K7_HUMAN STANDARD; PRT; 606 AA.
ID M3K7_HUMAN
AC O43318; O43317; O43319;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-
DE activated kinase 1).
GN MAP3K7 OR TAK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RP TISSUE=Lung;
RC MEDLINE=98153801; Pubmed=9480845;
RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.;
RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
RT NF-kappa B-inducing kinase-independent mechanism.";
RT Biochem. Biophys. Res. Commun. 243:545-549(1998).
RL -I- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.
CC -I- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1A, 1B (SHOWN HERE) AND 1C; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -I- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB009357; BAA25026.1; -
CC EMBL; AB009356; BAA25025.1; -.

```



DR EMBL; AB009358; BAA25027.2; -.  
 DR HSSP; P08631; 1AD5.  
 DR Genew; HGNC:6859; MAP3K7.  
 DR MIM; 602614; -.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR Transferase; Serine/threonine-protein kinase; ATP-binding;  
 KW Alternative splicing.  
 FT DOMAIN 8 14 POLY-SER.  
 FT NP\_BIND 36 291 PROTEIN\_KINASE.  
 FT BINDING 42 50 ATP (BY SIMILARITY).  
 FT ACT\_SITE 63 63 ATP (BY SIMILARITY).  
 FT ACT\_SITE 156 156 BY SIMILARITY.  
 FT VARSPPLIC 404 430 MISSING (IN ISOFORM 1A).  
 FT VARSPPLIC 509 518 PLAPCPNSKE -> ARTSCRTGPG (IN ISOFORM 1C).  
 FT VARSPPLIC 519 606 MISSING (IN ISOFORM 1C).  
 SQ SEQUENCE 606 AA; 67196 MW; 3DBF8147CD174013 CRC64;

Query Match 19.9%; Score 471.5; DB 1; Length 606;  
 Best Local Similarity 29.7%; Pred. No. 1.2e-23;  
 Matches 134; Conservative 82; Mismatches 150; Indels 85; Gaps 20;

QY 7 SFVQIKFDLQFFENCSSGSGSVYRAKMSIQDKEAVKLLKIEKAE-----IL 57  
 Db 27 NFEEDYKEIEVEEVGRGAFGVCAKAK--RAKDAIK--QIESESEKAFIVELROL 81  
 QY 58 SVLSHRNIIQFYGVILEPNNYGVTEYASLSGLDYINSNRSEEM---DMDHMTATDV 114  
 Db 82 SRVNHPIVLYGACLN--VCLVMEYAEAGSLYNVLHG--AEPLPYTTAAHAMSQCLQC 137  
 QY 115 AKGMHYLHMEAPVKVIHRDLKSRNVIAADG-VLKICDFG-ASRFHNHTHMSLVGTFPW 172  
 Db 138 SGGVAVLHSMQPKALIHRLKPPNLLVAGGTVLKICDFGTACDIQTHMTNNK--GSAAW 195  
 QY 173 MAPEVQLSLVSETCDTYSYGVVLWMLTREVPFKGLEG--LQVAVLVEKNERLTIPSS 230  
 Db 196 MAPEVFEGSNYSEKCDVFSWGIILWEVITRRKPFDEIGPAFRIMW-AVHNGTRPPLIKN 254  
 QY 231 CPRSFABLLHQCEADAKKRPSPKQIISIL-----ESMSNDT- 267  
 Db 255 LKPIESLMTRCWSKDPSPQSPMEIIVKIMTHLMRYFPGADEPLQYPCQYSDEGQNSAT 314  
 QY 268 ---SLPDKNSFLHNKAWECE-IEATLERLKKLERDLSFKEQELKERERRLKMEQKLT 323  
 Db 315 STGSFMDIASTNTSNKSDPTNMEQVPATNDTIKRLSKLLKNQAKQSESGRLSGASRGS 374  
 QY 324 EQSNTPLLLPLAARMSEESYFESKTESNSAEMSCQITATS---NGEGHGNPSLQAMM 379  
 Db 375 SVESLP---PTS-----EGKMSADMSIEARIAATTAYSKPKRGHRTAS----- 417  
 QY 380 LMFGGDIFSMNKAGAVMHSGMGINMOAKQNS 410  
 Db 418 ---FGNILDVPE---IVTSG---NGQPRRRS 439

RESULT 9  
 ABL\_MLVAB  
 ID\_ABL\_MLVAB STANDARD; PRT; 746 AA.  
 AC P00521;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tyrosine-protein kinase transforming protein ABL (EC 2.7.1.112).  
 OS V-ABL.  
 OS Abelson murine leukemia virus.  
 OC Viruses; Retrovirdae; Retroviridae; Mammalian type C retroviruses.

OX NCBI\_TaxID=11788;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83221648; PubMed=6304726;  
 RA Reddy E.P., Smith M.J., Srinivasan A.;  
 RT "Nucleotide sequence of Abelson murine leukemia virus genome:  
 RT structural similarity of its transforming gene product to other onc  
 RT gene products with tyrosine-specific kinase activity.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:3623-3627(1983).  
 RN (2)  
 RP REVISIONS TO 588-746.  
 RA Reddy E.P., Smith M.J., Srinivasan A.;  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:7372-7372(1983).  
 RN (3)  
 RP SEQUENCE OF 233-327 FROM N.A.  
 RX MEDLINE=83245023; PubMed=6191223;  
 RA Groffen J., Helstekamp N., Reynolds F.H. Jr., Stephenson J.R.;  
 RT "Homology between phosphotyrosine acceptor site of human c-abl and  
 RT viral oncogene products.";  
 RL Nature 304:167-169(1983).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL  
 CC POLYPROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; V01541; -; NOT ANNOTATED\_CDS.  
 DR EMBL; K00010; AAA46470.1; -.  
 DR PIR; A00627; TYWVGM.  
 DR HSSP; P00519; 2ABL.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR ProDom; PD000093; SH2; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR SMART; SM00219; TYRKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 KW Polyprotein; Tyrosine-protein kinase; Transferase; Oncogene;  
 KW SH2 domain.  
 FT DOMAIN 13 103 SH2.  
 FT DOMAIN 128 379 PROTEIN\_KINASE.  
 SQ SEQUENCE 746 AA; 81872 MW; B9072FFB5FE9257 CRC64;

Query Match 18.3%; Score 434; DB 1; Length 746;  
 Best Local Similarity 28.7%; Pred. No. 4.4e-21;  
 Matches 121; Conservative 73; Mismatches 172; Indels 56; Gaps 13;

QY 15 DLOFFENCSSGSGSVYRAKMSIQDKEAVKLL---LKIE---KEAEISVLSHRNIIQ 67  
 Db 127 DITMKHKLGGGQGYEGVWKKYSLTVAVKTLKEDTMEVEEFLKEAAVKEIKHPNLVQ 186  
 QY 68 FYGVILEPNNYGVTEYASLSGLDYINSNRSEEMDMDHMTATDVAKGMHYLHMEAPV 127  
 Db 187 LIGVCTREPPFYIITFMTYGNLLDYLRCNQGEVSAVLLYMATQISSAMEYLEKK--- 243





CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X16416; CAA34438.1; -  
DR EMBL; M14752; AAA51561.1; -  
DR EMBL; U07563; AAB60394.1; -  
DR EMBL; U07563; AAB60393.1; -  
DR EMBL; U07561; AAB60393.1; JOINED.  
DR EMBL; S69223; AAD14034.1; -  
DR PIR; A25582; TVHUA.  
DR PDB; 1AB2; 31-JAN-94.  
DR PDB; 1ABL; 01-NOV-94.  
DR PDB; 2ABL; 04-SEP-97.  
DR PDB; 1AWO; 28-JAN-98.  
DR PDB; 1BBZ; 25-NOV-98.  
DR Genew; HGNC:76; ABL1.  
DR MIM; 189980; -  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00017; SH2; 1.  
DR Pfam; PF00018; SH3; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00401; SH2DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR ProDom; PD000066; SH3; 1.  
DR ProDom; PD000093; SH2; 1.  
DR SMART; SM00252; SH2; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00219; TYRK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS50001; SH2; 1.  
DR PROSITE; PS50002; SH3; 1.  
KW Transferase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;  
KW Phosphorylation; SH2 domain; SH3 domain; Chromosomal translocation;  
KW 3D-structure; Alternative splicing.  
FT DOMAIN 61 121 SH3.  
FT DOMAIN 127 217 SH2.  
FT DOMAIN 242 493 PROTEIN KINASE.  
FT DOMAIN 605 609 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
FT DOMAIN 18 22 POLY-SER.  
FT DOMAIN 605 609 POLY-LYS.  
FT DOMAIN 782 1019 PRO-RICH.  
FT DOMAIN 897 903 POLY-PRO.  
FT SITE 26 27 BREAKPOINT FOR TRANSLOCATION TO FORM  
FT NP\_BIND 248 256 ATP (BY SIMILARITY).  
FT BINDING 271 271 ATP (BY SIMILARITY).  
FT ACT\_SITE 363 363 BY SIMILARITY.  
FT MOD\_RES 393 393 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT VARSPLIC 1 26 MLEICLVGCKSKGLSSSSCYCLE -> MGQPGKVLGD  
FT QRRPSLPALHPTKGAGKKESSRRHGPGHNVFEH (IN  
FT ISOFORM IB).  
FT CONFLICT 140 140 L -> P (IN REF. 2).  
FT CONFLICT 159 159 G -> S (IN REF. 2).  
FT CONFLICT 424 425 AF -> GK (IN REF. 4).  
FT CONFLICT 445 445 L -> R (IN REF. 2).  
FT CONFLICT 459 459 E -> K (IN REF. 2).  
FT CONFLICT 520 520 S -> T (IN REF. 2).  
FT CONFLICT 719 719 A -> V (IN REF. 2).  
FT CONFLICT 837 837 E -> G (IN REF. 3 AND 6).  
FT CONFLICT 837 837 E -> W (IN REF. 2).  
FT CONFLICT 863 863 G -> R (IN REF. 2).  
FT CONFLICT 894 894 R -> K (IN REF. 2).

FT CONFLICT 917 919 SPS -> RPG (IN REF. 2).  
FT CONFLICT 952 952 G -> A (IN REF. 2).  
FT CONFLICT 967 968 OS -> HP (IN REF. 2).  
FT CONFLICT 983 983 S -> LS (IN REF. 2).  
FT CONFLICT 1022 1022 MISSING (IN REF. 2).  
FT CONFLICT 1045 1045 R -> G (IN REF. 2).  
FT CONFLICT 1103 1103 T -> S (IN REF. 2).  
FT TURN 120 120

Query Match 17.3%; Score 410; DB 1; Length 1130;  
Best Local Similarity 25.9%; Pred. No. 2.7e-19;  
Matches 119; Conservative 80; Mismatches 183; Indels 78; Gaps 15;

QY 15 DLQFRENCGGSGFSGVYRAKMSQDKVAVKL-----LKIE--KEAELSVLSHRNIQ 67  
Db 241 DITMKHKLGGGQYGEVYEGVWKVSLTVAVKTLKEDTMEVEEFLKEAAVMKEIKHPNLVQ 300  
QY 68 FYGVILEPPNYGIVTEYASLSLYDINSNRSEMDMDHMTWATDVAKGMHYLMEAPV 127  
Db 301 LIGVCTREPPFYITTEFTYGNLLDYLRBCNROEVNAVLLYMATQISSANEYLEKK--- 357  
QY 128 KVIHRDLKSRNVIAADGVLCIDFGASRFHNTHTMSLVGT-FP--WMAPEVIQSLPVS 184  
Db 358 NFIRHDLAARNCLVGENHLVKVADFGLSRLMTGDTYTAHAGAKFPIKWTAPESLAYNKS 417  
QY 185 ETCDDTYSYGVVLMEMLTREV-PFKGLEGLQVAWLVEKNERLTIPSSGCRSFAELLHQW 243  
Db 418 IKSDVWAFGVLLWEIATYGMSPYPGIDLSQV-YELLEKDYRMERPEGPEKVELMRACW 476  
QY 244 EADAKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKLRLDSF 303  
Db 477 QWNPSDRPSFAEIHQAFETMFQESSISDEVEKELGKQ-----VRGAVSTL-----LQA 525  
QY 304 KEQELKERERRLKWEQKLTQSNTPLLPLAARMSESYFESKTE-ESNSAE---MSC 358  
Db 526 PELPTKTRTSRAAEHRDITDVPMP-----HSKQGESDPLDHEPAVSP 570  
QY 359 QITATNGEGHGNMPSLQAMMLMGFDIFSMNKAGAVMHSMQIMN-----QAKONSSKT 413  
Db 571 LIPRKERGPPEG-----GLNEDERLLPKOKTNLFSALIKKKKTAPT 613  
QY 414 TSKRRGKVNMALGSDFDLSEGDHDDDDDEEDNDMDN 453  
Db 614 PKRS-----SSFREMDCGPERRGAGEEGRDISN 643

RESULT 11  
ABLI\_MOUSE STANDARD; PRT; 1123 AA.  
ID ABLI\_MOUSE  
AC P00520; Q61252; Q61253; Q61254; Q61255; Q61256; Q61257; Q61258;  
AC Q61259; Q61260; Q61261; P97896;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Proto-oncogene tyrosine-protein kinase ABL1 (EC 2.7.1.112) (p150)  
DE (C-ABL).  
GN ABL1 OR ABL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=88068561; PubMed=3317402;  
RA Oppl C., Shore S.K., Reddy E.P.;  
RT "Nucleotide sequence of testis-derived c-abl cDNAs: implications for  
RL testis-specific transcription and abl oncogene activation.";  
RN Proc. Natl. Acad. Sci. U.S.A. 84:8200-8204(1987).  
RP SEQUENCE OF 1-187 FROM N.A. (ISOFORMS I TO IV).  
RX MEDLINE=95394474; PubMed=7665185;  
RA Chisoe S.L., Bodenteich A., Wang Y.-F., Wang Y.-P., Burian D.,

RA Clifton S.W., Crabtree J., Freeman A., Iyer K., Jian L., Ma Y.,  
RA McLaury H.-J., Pan H.-Q., Sarhan O.H., Toth S., Wang Z., Zhang G.,  
RA Helsterkamp N., Groffen J., Roe B.A.;  
RT "Sequence and analysis of the human ABL gene, the BCR gene, and  
RT regions involved in the Philadelphia chromosomal translocation";  
RL Genomics 27:67-82(1995).  
RN [3]  
RP SEQUENCE OF 85-182 FROM N.A.  
RX MEDLINE#84106840; PubMed=6319018;  
RA Wang J.Y.J., Ledley F., Goff S., Lee R., Groner Y., Baltimore D.;  
RT "The mouse c-abl locus: molecular cloning and characterization.";  
RL Cell 36:349-356(1984).  
RN [4]  
RP ALTERNATIVE SPLICING.  
RX MEDLINE=88202920; PubMed=3283651;  
RA Bernards A., Paskind M., Baltimore D.;  
RT "Four murine c-abl mRNAs arise by usage of two transcriptional  
RT promoters and alternative splicing";  
RL Oncogene 2:297-304(1988).  
RN [5]  
RX X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 61-121.  
RA MEDLINE=95393198; PubMed=7664083;  
RT Musacchio A., Saraste M., Wilmanns M.;  
RT "High-resolution crystal structures of tyrosine kinase SH3 domains  
RT complexed with proline-rich peptides";  
RL Nat. Struct. Biol. 1:546-551(1994).  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; THE MYRISTOYLATED C-ABL PROTEIN  
CC WAS REPORTED TO BE NUCLEAR.  
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; I (SHOWN HERE), II, III AND  
CC IV; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
CC SUBFAMILY.  
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J02995; AAA88241.1; -;  
DR EMBL; U14721; AAB60451.1; -;  
DR EMBL; U14720; AAB60451.1; JOINED.  
DR EMBL; U14721; AAB60450.1; -;  
DR EMBL; U14720; AAB60450.1; JOINED.  
DR EMBL; U14721; AAB60448.1; -;  
DR EMBL; U13835; AAB60448.1; JOINED.  
DR EMBL; U14721; AAB60449.1; -;  
DR EMBL; U13835; AAB60449.1; JOINED.  
DR EMBL; X07539; CAA30412.1; -;  
DR EMBL; X07540; CAA30413.1; -;  
DR EMBL; X07541; CAA30414.1; -;  
DR EMBL; M12263; AAA37136.1; -;  
DR EMBL; M12264; AAA37137.1; -;  
DR EMBL; M12265; AAA37138.1; -;  
DR EMBL; M12266; AAA37134.1; -;  
DR EMBL; K03228; AAA37135.1; -;  
DR PIR; A00626; A00626.  
DR PIR; A39962; A39962.  
DR PDB; LABO; 15-OCT-95.  
DR PDB; MGT; 87859; ABl1.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_pkinase

| Accession | Protein Name   | Length | Start      | End | Score | Model | Annotations                              |
|-----------|--|--------|------------|-----|-------|-------|--|
| DR        | Pfam; PF00017; SH2; 1.   | 184    | 187        |     |       |       |  |
| DR        | Pfam; PF00018; SH3; 1.   | 1123   | AA; 122676 |     |       |       |  |
| DR        | Pfam; PF00069; kinase; 1.  |        |            |     |       |       |  |
| DR        | PRINTS; PR00401; SH2DOMAIN.  |        |            |     |       |       |  |
| DR        | PRINTS; PR00109; TYRKINASE.  |        |            |     |       |       |  |
| DR        | ProDom; PD000001; Euk_kinase; 1.                                   |        |            |     |       |       |  |
| DR        | ProDom; PD000066; SH3; 1.  |        |            |     |       |       |  |
| DR        | ProDom; PD000093; SH2; 1.  |        |            |     |       |       |  |
| DR        | PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.                           |        |            |     |       |       |  |
| DR        | PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.                           |        |            |     |       |       |  |
| DR        | PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.                           |        |            |     |       |       |  |
| DR        | PROSITE; PS50001; SH2; 1.  |        |            |     |       |       |  |
| DR        | PROSITE; PS50002; SH3; 1.  |        |            |     |       |       |  |
| KW        | Transferase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding; |        |            |     |       |       |  |
| KW        | phosphorylation; SH2 domain; SH3 domain; Alternative splicing;     |        |            |     |       |       |  |
| KW        | Nuclear protein; 3D-structure.                                     |        |            |     |       |       |  |
| FT        | DOMAIN   | 61     | 121        |     |       |       | SH3.                                     |
| FT        | DOMAIN   | 127    | 217        |     |       |       | SH2.                                     |
| FT        | DOMAIN   | 242    | 493        |     |       |       | PROTEIN KINASE.                          |
| FT        | DOMAIN   | 605    | 609        |     |       |       | NUCLEAR LOCALIZATION SIGNAL (POTENTIAL). |
| FT        | DOMAIN   | 18     | 22         |     |       |       | POLY-SER.                                |
| FT        | DOMAIN   | 605    | 609        |     |       |       | POLY-LYS.                                |
| FT        | DOMAIN   | 804    | 1012       |     |       |       | PRO-RICH.                                |
| FT        | DOMAIN   | 891    | 897        |     |       |       | POLY-PRO.                                |
| FT        | NP_BIND  | 248    | 256        |     |       |       | ATP (BY SIMILARITY).                     |
| FT        | ACT_SITE   | 271    | 271        |     |       |       | ATP (BY SIMILARITY).                     |
| FT        | MOD_RES  | 363    | 363        |     |       |       | BY SIMILARITY.                           |
| FT        | VARSPPLIC  | 1      | 26         |     |       |       | PHOSPHORYLATION (AUTO-) (BY SIMILARITY). |
| FT        | VARSPPLIC  | 1      | 26         |     |       |       | MLEICLKVCGSKSGLSSSSSCYCLE -> MSQRWYTKCR  |
| FT        | VARSPPLIC  | 1      | 26         |     |       |       | VQDPALPFM (IN ISOFORM III).              |
| FT        | VARSPPLIC  | 1      | 26         |     |       |       | MLEICLKVCGSKSGLSSSSSCYCLE -> MISFDLSDDEL |
| FT        | VARSPPLIC  | 1      | 26         |     |       |       | HKLVLVDV (IN ISOFORM II).                |
| FT        | VARSPPLIC  | 1      | 26         |     |       |       | MLEICLKVCGSKSGLSSSSSCYCLE -> MGQDPKVLGD  |
| FT        | VARSPPLIC  | 1      | 26         |     |       |       | ORPSPALHFIKAGKRDSSRHGGPHCNVFEH (IN       |
| FT        | VARSPPLIC  | 1      | 26         |     |       |       | ISOFORM IV).                             |
| FT        | VARSPPLIC  | 1      | 26         |     |       |       | LYVS -> VGDW (IN REF. 2).                |
| FT        | VARSPPLIC  | 1      | 26         |     |       |       | MM; 284F0830644AFD8F CRC64;              |

| Query Match           | 17.1%;       | Score 405;              | DB 1;                    | Length 1123;               |
|-----------------------|--------------|-------------------------|--------------------------|----------------------------|
| Best Local Similarity | 27.2%;       | Pred. No. 5.7e-19;      |                          |                            |
| Matches 125;          | Conservative | 78;                     | Mismatches 179;          | Indels 78;                 |
|                       |              |                         |                          | Gaps 17;                   |
| Qy                    | 15           | DLQFFENC GGSGFSGSVYRAK  | WISQDKEVAVKKL---         | LKIE---KEAEILSVLSHRNIIQ 67 |
|                       |              | :                       |                          | :::     : :     :          |
| Db                    | 241          | DITMKHKLGGGQYGEVYEGVAKK | YSLTVAVKTLKEDTMEVEEFLKEA | AVMKEIKHPNLVQ 300          |
| Qy                    | 68           | FYGVILEPPNYGIVTEYASLSL  | YDIYNSNRSEEMDMHINTWATD   | VAKGHHYLMHAPV 127          |
|                       |              | : : : : : : : : : :     | : : : : : : : :          | : : : : : : :              |
| Db                    | 301          | ILGVCTREPPFYIITEFMTYGN  | LIDYLRBCNRQEVSAVLLYMATQ  | ISSAMEYLEKK--- 357         |
| Qy                    | 128          | KVIHRDLKSRNVIAADGVLCI   | DEGASRFHNHTHMSLVGT-PP-   | WMAPEVIQSLPVS 184          |
|                       |              | : : : : : : : :         | : : : : : : : :          | : : : : : : :              |
| Db                    | 358          | NEIHRDLARNCLVGENHLVKAD  | FGLSRLMTGDTYTAHAGAKFP    | IKWTAPESLAYNKF 417         |
| Qy                    | 185          | ETCDTYSYGVLMEMLTREV-    | PEKGLEGLQVAWLVEKENERLT   | IPSSCPSPFAELHQW 243        |
|                       |              | : : : : : : : : : :     | : : : : : : : : : :      | : : : : : : : : :          |
| Db                    | 418          | IKSDVMAFGVLLMEIATYGM    | SFPYGLDLSQV-YELLEKDYR    | MERPEGCEKVELMRACW 476      |
| Qy                    | 244          | EADAKRPPSFKOIISILES     | MSNDTSLPDKCNSFLHNKA      | EWRCIEATLERLKLERD--- 300   |
|                       |              | : : : : : : : : : :     | : : : : : : : : : :      | : : : : : : : : : :        |
| Db                    | 477          | QWNPSDRPSPFAIHOAFET     | MFQESSISD-----           | EVEKELG--KRGTREGAG 520     |
| Qy                    | 301          | --LSFKEQLKERERRLK       | WMEQKLTQESNTPLLPLAAR     | MSSESYFESKTESSNSAEMSC 358  |
|                       |              | :   :   :   :   :   :   | :   :   :   :   :   :    | :   :   :   :   :   :      |
| Db                    | 521          | SMLQAPLPTKTRTCR-RAE     | QK--DADTPELL-----        | HTKGLGESDALDSEPA-VSP 570   |
| Qy                    | 359          | QITATSNGBHGMPSLQAM      | MLMGFDIFS MNKAGAVM       | HSGMOINM-----QAKONSSKT 413 |
|                       |              | : : : : : : : : : :     | : : : : : : : : : :      | : : : : : : : : : :        |
| Db                    | 571          | LIPRKERGPDPG-----       | SINEDERLLPRDKTINL        | FSALIKKKKKMAPT 613         |
| Qy                    | 414          | TSKRRGKKVNMALGFS        | DFDLSEGD DDDDDGEEED      | NMDMN 453                  |
|                       |              | :   : : : : : : :       | : : : : : : : :          | : : : : : : : :            |
| Db                    | 614          | PPKRS-----SSF           | REMDGQPD RRGASEDD        | SREL CN 643                |

```

RESULT 12
ABL2_HUMAN          STANDARD;          PRT; 1182 AA.
ID   ABL2_HUMAN
AC   P42684;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Tyrosine-protein kinase ABL2 (EC 2.7.1.112) (Tyrosine kinase ARG).
GN   ABL2 OR ARG OR ABL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90332670; PubMed=2198571;
RA   Krüh G.D., Perego R., Miki T., Aaronson S.A.;
RT   "The complete coding sequence of arg defines the Abelson subfamily of
RT   cytoplasmic tyrosine kinases."
RL   Proc. Natl. Acad. Sci. U.S.A. 87:5802-5806(1990).
CC   -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC   tyrosine phosphate.
CC   -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A AND 1B (SHOWN HERE); ARE
CC   PRODUCED BY ALTERNATIVE SPLICING.
CC   -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC   -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
CC   SUBFAMILY.
CC   -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC   -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC   -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC   WWW="http://www.infobiogen.fr/services/chromancer/Genes/ABL2ID226.html".
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; M35296; AAA35553.1; -.
DR   HSSP; P00519; 1BB2.
DR   Genew; HGNC:77; ABL2.
DR   MIM; 164690; -.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001452; SH3.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF00018; SH3; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00401; SH2DOMAIN.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Euk_pkinase; 1.
DR   ProDom; PD000066; SH3; 1.
DR   ProDom; PD000093; SH2; 1.
DR   SMART; SM00252; SH2; 1.
DR   SMART; SM00326; SH3; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50001; SH2; 1.
DR   PROSITE; PS50002; SH3; 1.
DR   Transferrase; Tyrosine-protein kinase, Proto-oncogene, ATP-binding;
KW   Phosphorylation; SH2 domain; SH3 domain; Alternative splicing.
FT   DOMAIN 107 167
FT   DOMAIN 173 263
FT   DOMAIN 288 539
FT   DOMAIN 561 564
FT   POLY-SER.

```

```

FT   DOMAIN 658 660
FT   DOMAIN 732 739
FT   DOMAIN 843 1055
FT   DOMAIN 984 988
FT   NP BIND 294 302
FT   BINDING 317 317
FT   ACT_SITE 409 409
FT   MOD_RES 439 439
FT   VARSPLIC 1 73
FT
FT
SQ   SEQUENCE 1182 AA; 128343 MW; ED93869BC2B14FAA CRC64;
Query Match 17.1%; Score 405; DB 1; Length 1182;
Best Local Similarity 32.8%; Pred. No. 6.1e-19;
Matches 113; Conservative 59; Mismatches 144; Indels 28; Gaps 13;
QY 5 GASFVQIKFD---DLQFFENCGGSGFSGVYRAKNIQDKEVAVKCL---LKIE---KE 53
Db 273 GVSPIDHKWEMERTDITMKHKLGGQYGEVYGVWKKYSLTVAVKTLKEDTMEVEEFLKE 332
QY 54 AEILSVLSHRNIIQFYGV-ILEPPNYGIYTEYASLSLYDINSNRSEMDMHIMTWAT 112
Db 333 AAVMKEIKHPNLVQLGCTLEPPFY-ITETVMPYGNLLDYLRECNREEVTAVVLLYMAT 391
QY 113 DVAKGMHYLHMEAPVYVIHRDLKSRNVIAADGVLCIDFGASRFHNTHTMSLVGT-PP 171
Db 392 QISSAMEYLEKK---NFIHRDLAARNCLVGENHYVAVDFGLSRMLTGDTYTAHAKAFP 448
QY 172 --WMAPEVIQSLPVSETCDTYSYGVVLMELTREV-PFKGLEGLQVAVLVEKNERLTIP 228
Db 449 IKWTAPESLAYNTSPISKSDVWAFGLVLTWEIATYGMSPYPGIDLSQV-YDLEKGYRMEQP 507
QY 229 SSCPRSFAELLHQCWEADAKRPSFKQIISLESMSNDTSLPDKCNSEFLHKAWEKCEIE 288
Db 508 ECGPPKVELMRAQWCKWSPADRPSPFAETHQAFETMFHDSISEVAEEL-GRAASSSSVV 566
QY 289 ATLERLKKLERDISFKEQLKERERRLKWE--QKLTQSNTP 330
Db 567 PYLRLPLPLPS---KTRTLKKOVENKENTEGAQDATENSASSL 606
RESULT 13
ABL_DROME          STANDARD;          PRT; 1520 AA.
ID   ABL_DROME
AC   P00522;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Tyrosine-protein kinase Abl (EC 2.7.1.112) (D-abn).
GN   ABL OR DASH OR ABL-1.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88174728; PubMed=2832740;
RA   Henkemeyer M.J., Bennett R.L., Gertler F.B., Hoffmann F.M.;
RT   "DNA sequence, structure, and tyrosine kinase activity of the
RT   Drosophila melanogaster Abelson proto-oncogene homolog."
RL   Mol. Cell. Biol. 8:843-853(1988).
RN   [2]
RP   SEQUENCE OF 374-648 FROM N.A.
RX   MEDLINE=84082064; PubMed=6317185;
RA   Hoffmann F.M., Fresco L.D., Hoffman-Falk H., Shilo B.-Z.;
RT   "Nucleotide sequences of the Drosophila src and abl homologs:
RT   conservation and variability in the src family oncogenes."
RN   [3]
RP   FUNCTION.

```





DR SMART; SM00454; SAM; 1.  
DR SMART; SM00449; SPY; 3.  
DR SMART; SM00221; STYK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS50105; SAM\_DOMAIN; 1.  
KW Transferase; Tyrosine-protein kinase; ATP-binding; phosphorylation.  
FT DOMAIN 908 972 SAM.  
FT DOMAIN 403 420 POLY-ASN.  
FT DOMAIN 428 435 POLY-THR.  
FT DOMAIN 449 480 POLY-ASN.  
FT DOMAIN 483 491 POLY-ASN.  
FT DOMAIN 494 508 POLY-ASN.  
FT DOMAIN 512 532 POLY-ASN.  
FT DOMAIN 596 600 POLY-ASN.  
FT DOMAIN 808 811 POLY-PHE.  
FT DOMAIN 1026 1029 POLY-SER.  
FT DOMAIN 1195 1210 POLY-ASN.  
FT DOMAIN 1215 1220 POLY-GLN.  
FT DOMAIN 1224 1233 POLY-GLN.  
FT DOMAIN 1266 1274 POLY-PRO.  
FT DOMAIN 1289 1561 PROTEIN\_KINASE.  
FT NP\_BIND 1295 1303 ATP (BY SIMILARITY).  
FT BINDING 1316 1316 ATP (BY SIMILARITY).  
FT ACT\_SITE 1417 1417 BY SIMILARITY.  
FT CONFLICT 1248 1248 D -> R (IN REF. 2).  
FT CONFLICT 1435 1435 V -> L (IN REF. 2).  
SQ SEQUENCE 1584 AA; 174304 MW; 5D1589458D8E01E3 CRC64;

Query Match 16.9%; Score 401.5; DB 1; Length 1584;  
Best Local Similarity 34.6%; Pred. No. 1.5e-18;  
Matches 102; Conservative 52; Mismatches 104; Indels 37; Gaps 8;

QY 10 QIKRDDLQFFENCSSGSGSVYRAKWSIDKEAVK-----KLKIEKEAEITLS 58  
Db 1283 EIDFNELEFGQTIGKGFGEVKGWY--RETDAIKIYRDQFKTKSLVMFQNEVGILS 1340  
QY 59 VLSHRNIIQFYGVILE--PPNYGIVTEYASLSGLDYINSNRSEEMDMHI--MTNATDVA 115  
Db 1341 KLRHBNVQFLGACTAGGEDHCHIVTEWGGSLRQFLTHFNLLLEQNPHIRLKLALDIA 1400  
QY 116 KGMHYLHMEAPVKIHRDLKSRNVI-----AADGVLKICDFGASRFHNTT 162  
Db 1401 KGMHYLHGWTP-FILHRDLSSGNILLDNIDPKNPVSSRODICKKISDFGLSLKKEQA 1459  
QY 163 HM--SLVGFPMWAPVIOGLVSETCDTYSYGVVLTWMLTREVPFKGLEGLQVAVLVE 220  
1460 SQMTQSVGCIPTMAPEVFKGDSNSEKSDVSYGMVLFELLTSDPEQDMKPMKMAHLAAY 1519  
QY 221 KNERLTISSCPSPFAELHOCWEADAKRPSFKQITISLES-----SNDTSLP 270  
Db 1520 ESYRPPPLTSSKWKELTQCWDSNPDSRPTFKQITVHLKEMEDQGVSPFASVP 1574

RESULT 15  
ABL\_FSVHY STANDARD; PRT; 439 AA.  
ID ABL\_FSVHY  
AC P10447;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Tyrosine-protein kinase transforming protein ABL (EC 2.7.1.112).  
GN V-ABL.  
OS Feline sarcoma virus (strain Hardy-Zuckerman 2).  
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus.  
OX NCBI\_TaxID=11776;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87141338; PubMed=3029415;  
RA Bergold P.J., Blumenthal J.A., D'Andrea E., Snyder H.W. Jr.,  
RA Lederman L., Silverstone A., Nguyen H., Besmer P.;  
RT "Nucleic acid sequence and oncogenic properties of the HZ2 feline

RT sarcoma virus v-abl insert."; J. Virol. 61:1193-1202(1987).  
CC 1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.  
CC 1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL-POLYPROTEIN.  
CC 1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL SUBFAMILY..  
CC 1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
CC 1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M15805; AAA43042.1; --  
DR PIR; A26132; TVMVAB.  
DR HSSP; P00519; 1BBZ.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00017; SH2; 1.  
DR Pfam; PF00018; SH3; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00401; SH2DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR ProDom; PD000066; SH3; 1.  
DR ProDom; PD000093; SH2; 1.  
DR SMART; SM00252; SH2; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00219; TYRK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS50002; SH3; 1.  
KW SH2 domain; SH3 domain.  
KW SH2 domain; Tyrosine-protein kinase; Transferase; Oncogene;  
FT DOMAIN 10 70 SH3.  
FT DOMAIN 76 166 SH2.  
FT DOMAIN 191 439 PROTEIN\_KINASE.  
SQ SEQUENCE 439 AA; 50004 MW; 135799EDFED1481AB CRC64;

Query Match 16.2%; Score 385.5; DB 1; Length 439;  
Best Local Similarity 33.2%; Pred. No. 3.3e-18;  
Matches 84; Conservative 53; Mismatches 101; Indels 15; Gaps 7;

QY 15 DLQFFENCSSGSGSVYRAKWSIDKEAVK-----LKIE---KEAELISVLSHRNIIQ 67  
Db 190 DITMKHKLGGQYGEVYEGVWKKYSLTVAVKTLKEDTMEVEERFLKEAAVMEIKHPNLVQ 249  
QY 68 FYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPV 127  
Db 250 LIGVCTREPPFYITTEFMTYGNLDYLRCCNRQEVNAVVLVYMATOISSAMEYLEKK--- 306  
QY 128 KVIHRDLKSRNVIAADGVLCDFGASRFHNTTHMSLVGT-FP--WMAPEVIOGLPVS 184  
Db 307 NFIRHDLAARNCLVGENHLVYADDFGLSLRLMTGDTYTAAGTKPKWTAPESLAYNKS 366  
QY 185 ETCPTYSYGVVLTWMLTREVPFKGLEGLQVAVLVEKNERLTISSCPSPFAELHOCW 243  
Db 367 IKSDVWAFGLVLMETATYGMSPYIGIDLSQV-YELLEKDYRMERPEGCEPEKVEELMRAW 425  
QY 244 EADAKRPSFKQI 256  
Db 426 QWPSDRPAFAEI 438

Sun May 4 10:01:18 2003

us-09-757-982-5.rsp

Page 14

Search completed: May 1, 2003, 20:36:17  
Job time : 28.8452 secs

---

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:28:47 ; Search time 50.7671 Seconds  
(without alignments)  
1846.697 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 2375

Sequence: 1 MSSLGASFVQIKFDLDQFFE.....GDDDDDDGEEEDNDMDNSE 455

Scoring table:

BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

al number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriophage:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID     | Description        |
|------------|--------|-------------|--------|-----------|--------------------|
| 1          | 2375   | 100.0       | 455    | 4 Q9HCC4  | Q9hcc4 homo sapien |
| 2          | 2299.5 | 96.8        | 454    | 11 Q9ESL3 | Q9esl3 mus musculu |
| 3          | 1751.5 | 73.7        | 800    | 4 Q9NYL2  | Q9nyl2 homo sapien |
| 4          | 1749.5 | 73.7        | 800    | 4 Q9HCC5  | Q9hcc5 homo sapien |
| 5          | 1748.5 | 73.6        | 800    | 4 Q9HDD2  | Q9hdd2 homo sapien |
| 6          | 1748.5 | 73.6        | 800    | 4 Q9NVE9  | Q9nve9 homo sapien |
| 7          | 1743.5 | 73.4        | 802    | 11 Q9ESL4 | Q9esl4 mus musculu |
| 8          | 1402.5 | 59.1        | 371    | 13 Q90ZY8 | Q90zy8 brachydanio |
| 9          | 601    | 25.3        | 1066   | 4 Q9H2N5  | Q9h2n5 mus musculu |
| 10         | 598.5  | 25.2        | 1001   | 11 Q8VDG6 | Q8vdg6 mus musculu |
| 11         | 597.5  | 25.2        | 564    | 4 Q9HLY7  | Q9hly7 homo sapien |
| 12         | 590.5  | 24.9        | 1036   | 4 Q8WMN1  | Q8wmn1 homo sapien |
| 13         | 588.5  | 24.8        | 570    | 4 Q8WMN2  | Q8wmn2 homo sapien |
| 14         | 581.5  | 24.5        | 847    | 4 Q16584  | Q16584 homo sapien |
| 15         | 580.5  | 24.4        | 850    | 11 Q9JUI5 | Q9jui5 mus musculu |
| 16         | 572.5  | 24.1        | 740    | 5 Q21982  | Q21982 caenorhabdi |

|    |       |      |      |           |                    |
|----|-------|------|------|-----------|--------------------|
| 17 | 568.5 | 23.9 | 966  | 4 Q43283  | Q43283 homo sapien |
| 18 | 565.5 | 23.8 | 859  | 4 Q8WY25  | Q8wy25 homo sapien |
| 19 | 560.5 | 23.6 | 406  | 10 Q23719 | Q23719 arabidopsis |
| 20 | 560.5 | 23.6 | 880  | 10 Q8S9K4 | Q8s9k4 arabidopsis |
| 21 | 544.5 | 22.9 | 1020 | 5 Q9W3I3  | Q9w3i3 drosophila  |
| 22 | 544.5 | 22.9 | 1148 | 5 Q95VF6  | Q95vf6 drosophila  |
| 23 | 544.5 | 22.9 | 1161 | 5 Q95UN8  | Q95un8 drosophila  |
| 24 | 529   | 22.3 | 886  | 10 Q9LYI8 | Q9lyi8 arabidopsis |
| 25 | 527.5 | 22.2 | 977  | 5 Q9VM24  | Q9vm24 drosophila  |
| 26 | 510.5 | 21.5 | 1030 | 10 Q9C9U5 | Q9c9u5 arabidopsis |
| 27 | 505   | 21.3 | 855  | 5 Q01700  | Q01700 caenorhabdi |
| 28 | 495   | 20.8 | 982  | 10 Q65833 | Q65833 lycopersico |
| 29 | 494   | 20.8 | 903  | 10 Q9FPR5 | Q9fpr5 oryza sativ |
| 30 | 483.5 | 20.4 | 847  | 10 Q93XL9 | Q93xl9 rosa hybrid |
| 31 | 483   | 20.3 | 491  | 4 Q9NZ70  | Q9nz70 homo sapien |
| 32 | 480   | 20.2 | 829  | 10 Q24027 | Q24027 lycopersico |
| 33 | 478.5 | 20.1 | 933  | 10 Q9FPR3 | Q9fpr3 arabidopsis |
| 34 | 475.5 | 20.0 | 412  | 10 Q9M085 | Q9m085 arabidopsis |
| 35 | 475   | 20.0 | 957  | 10 Q9FPR4 | Q9fpr4 hordeum vul |
| 36 | 472.5 | 19.9 | 616  | 13 Q73613 | Q73613 xenopus lae |
| 37 | 471.5 | 19.9 | 606  | 11 Q923A8 | Q923a8 mus musculu |
| 38 | 471   | 19.8 | 777  | 10 Q9C833 | Q9c833 arabidopsis |
| 39 | 471   | 19.8 | 806  | 10 Q9ZSD8 | Q9zsd8 lycopersico |
| 40 | 471   | 19.8 | 829  | 10 Q9ZSD9 | Q9zsd9 lycopersico |
| 41 | 470.5 | 19.8 | 763  | 10 Q9C903 | Q9c903 arabidopsis |
| 42 | 470   | 19.8 | 735  | 10 Q93ZU3 | Q93zu3 arabidopsis |
| 43 | 470   | 19.8 | 736  | 10 Q82754 | Q82754 arabidopsis |
| 44 | 468.5 | 19.7 | 678  | 5 Q9V3Q6  | Q9v3q6 drosophila  |
| 45 | 468   | 19.7 | 539  | 4 Q9NTRI  | Q9ntri homo sapien |

#### ALIGNMENTS

RESULT 1

Q9HCC4 PRELIMINARY; PRT; 455 AA.

AC Q9HCC4; AC Q9HCC4;

DT 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE MLTK-beta (Similar to sterile-alpha motif and leucine zipper containing kinase AZK) (Mixed lineage kinase) (Mixed lineage kinase-related kinase MRK-beta).

DE MLTK.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. NCBI\_TaxID=9606;

OX [1]

RP SEQUENCE FROM N.A.

RA Gotoh I., Adachi M., Nishida E.;

RT "Identification and Characterization of a Novel MAP Kinase Kinase Kinase, MLTK.";

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=COLON;

RA Strausberg R.;

| RL | Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases. |
| RN | [3] |
| RP | SEQUENCE FROM N.A. |
| RA | Acton S.; |
| RT | "MLK-mixed lineage kinase."; |
| RL | Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases. |
| RN | [4] |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE=21950776; PubMed=11836244; |
| RA | Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.; |
| RT | "MRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in gamma-Radiation-induced Cell Cycle Arrest."; |
| RL | J. Biol. Chem. 277:13873-13882(2002). |
| CC | -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. |

DR EMBL; AB049734; BAB16445.1; -  
DR EMBL; BC001401; AAH01401.1; -  
DR EMBL; AF325454; AAK11615.1; -  
DR EMBL; AF480462; AAL85892.1; -  
DR HSSP; P12931; 1FMK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00221; STYKc; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; TYRKc; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 455 AA; 51582 MW; E87DB84A4D58B752 CRC64;

Query Match 100.0%; Score 2375; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.1e-174;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDLQFFENC GGSGSVYRAKWSQDKKVAVKLLKIEKAEILSVL 60  
DB 1 MSSLGASFVQIKFDLQFFENC GGSGSVYRAKWSQDKKVAVKLLKIEKAEILSVL 60  
QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEMDMHMTWATDVAKGMHY 120  
DB 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEMDMHMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
DB 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
QY 181 LPVSETCDTYSYGVLMEMLTREVPFKGLEQLQVAMLVVEKNERLTIPSSCPRSFAELLH 240  
DB 181 LPVSETCDTYSYGVLMEMLTREVPFKGLEQLQVAMLVVEKNERLTIPSSCPRSFAELLH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKKLERD 300  
DB 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKKLERD 300  
QY 301 LSFKEQELKERERRLKMEQKLTQSNTPLLPLAARMESESYFESKTEESNSAEMSCQI 360  
DB 301 LSFKEQELKERERRLKMEQKLTQSNTPLLPLAARMESESYFESKTEESNSAEMSCQI 360  
QY 361 TATSNGEGHGMNPSLQAMLMGFGDIFS MNKAGAVMHSQMGINMOAKONSSKTTSKRRGK 420  
DB 361 TATSNGEGHGMNPSLQAMLMGFGDIFS MNKAGAVMHSQMGINMOAKONSSKTTSKRRGK 420  
QY 421 KVNMLGFSDFDLSEGDDDDDDGEEEDNDMDNSE 455  
DB 421 KVNMLGFSDFDLSEGDDDDDDGEEEDNDMDNSE 455

RESULT 2  
Q9ESL3  
ID Q9ESL3 PRELIMINARY; PRT; 454 AA.  
AC Q9ESL3;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE MLTK-beta.  
GN ZAK OR MLTK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE-FROM N.A.  
RX MEDLINE=2f264927; PubMed=11042189;

RA Gotch I.; Adachi M.; Nishida E.;  
RT "Identification and Characterization of a Novel MAP Kinase Kinase  
RT Kinase, MLTK.";  
RL J. Biol. Chem. 276:4276-4286(2001).  
CC - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AB049732; BAB16443.1; -  
DR HSSP; P12931; 1FMK.  
DR MGD; MGI:1931274; Zak.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00221; STYKc; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; TYRKc; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR ATP-binding; Serine/threonine-protein kinase; Transferase.  
KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 454 AA; 51366 MW; 35C2FC0D729D9395 CRC64;

Query Match 96.8%; Score 2299.5; DB 11; Length 454;  
Best Local Similarity 96.5%; Pred. No. 7.1e-169;  
Matches 439; Conservative 10; Mismatches 5; Indels 1; Gaps 1;

QY 1 MSSLGASFVQIKFDLQFFENC GGSGSVYRAKWSQDKKVAVKLLKIEKAEILSVL 60  
DB 1 MSSLGASFVQIKFDLQFFENC GGSGSVYRAKWSQDKKVAVKLLKIEKAEILSVL 60  
QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEMDMHMTWATDVAKGMHY 120  
DB 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEMDMHMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
DB 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
QY 181 LPVSETCDTYSYGVLMEMLTREVPFKGLEQLQVAMLVVEKNERLTIPSSCPRSFAELLH 240  
DB 181 LPVSETCDTYSYGVLMEMLTREVPFKGLEQLQVAMLVVEKNERLTIPSSCPRSFAELLH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKKLERD 300  
DB 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKKLERD 300  
QY 301 LSFKEQELKERERRLKMEQKLTQSNTPLLPLAARMESESYFESKTEESNSAEMSCQI 360  
DB 301 LSFKEQELKERERRLKMEQKLTQSNTPLLPLAARMESESYFESKTEESNSAEMSCQI 360  
QY 361 TATSNGEGHGMNPSLQAMLMGFGDIFS MNKAGAVMHSQMGINMOAKONSSKTTSKRRGK 420  
DB 361 TATSNGEGHGMNPSLQAMLMGFGDIFS MNKAGAVMHSQMGINMOAKONSSKTTCKRRGK 420  
QY 421 KVNMLGFSDFDLSEGDDDDDDGEEEDNDMDNSE 455  
DB 421 KVNMLGFSDFDLSEG-DDDDHDGDDAENDVDNSE 454

RESULT 3  
Q9NYL2  
ID Q9NYL2 PRELIMINARY; PRT; 800 AA.  
AC Q9NYL2;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Mixed lineage kinase.  
GN Homo sapiens (Human).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]



RP SEQUENCE FROM N.A.  
RX MEDLINE=20384179; PubMed=10924358;  
RA Liu T.C., Huang C.J., Chu Y.C., Wei C.C., Chou C.C., Chou M.Y.,  
RT Chou C.K., Yang J.J.;  
RT "Cloning and expression of ZAK, a mixed lineage kinase-like protein  
containing a leucine-zipper and a sterile-alpha motif."  
RL Biochem. Biophys. Res. Commun. 274:811-816(2000).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AF238255; AAF63490.1; -  
DR HSSP; P12931; 1FMK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR001660; SAM.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00536; SAM; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00454; SAM; 1.  
DR SMART; SM00221; STYKc; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 800 AA; 91264 MW; DA82D7AB2082F43 CRC64;

Query Match 73.7%; Score 1751.5; DB 4; Length 800;  
Best Local Similarity 76.1%; Pred. No. 2.4e-126;  
Matches 360; Conservative 24; Mismatches 56; Indels 33; Gaps 9;

QY 1 MSSLGASFVQIKFDDLOFFENC GGSGFSGSVYRAKWI SQDKEVA VKKLKIEKEA EILSVL 60  
Db 1 MSSLGASFVQIKFDDLOFFENC GGSGFSGSVYRAKWI SQDKEVA VKKLKIEKEA EILSVL 60  
QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDYINSNRSEEMDMDHIMTATDVAKGMHY 120  
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDYINSNRSEEMDMDHIMTATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180  
QY 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAVLVVEKNERLTIPSSCPRSFAELLH 240  
Db 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAVLVVEKNERLTIPSSCPRSFAELLH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIEATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIEATLERLKLKLERD 300  
QY 301 LSFKEQELKERERRLKMEQKLTQSNTPLL-LPLAARMSSESYF--ESKTEESNSAE 355  
Db 301 LSFKEQELKERERRLKMEQKLTQSNTPLL-LPLAARMSSESYF--ESKTEESNSAE 355  
QY 356 MSCQITATSNGECHGNPQLQAMLMFGDIFSMN--KAGAVMH--SGMQ-----INM- 404  
Db 356 MSCQITATSNGECHGNPQLQAMLMFGDIFSMN--KAGAVMH--SGMQ-----INM- 404  
QY 405 ---QAKONSSKTSKRGRKGVNMAFGSDFDLSGDDDD-----DGEE 446  
Db 405 ---QAKONSSKTSKRGRKGVNMAFGSDFDLSGDDDD-----DGEE 446  
QY 416 HFPPLIKDSGGEPEENEKIIVLELVFG-FHLKPGTGPDCKWKMYMEMDGE 467  
Db 416 HFPPLIKDSGGEPEENEKIIVLELVFG-FHLKPGTGPDCKWKMYMEMDGE 467

RESULT 4  
Q9HCCS PRELIMINARY; PRT; 800 AA.  
AC Q9HCCS;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DE MLTK-alpha. 20, Last annotation update)  
GN MLTK.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21264927; PubMed=11042189;  
RA Gotoh I., Adachi M., Nishida E.;  
RT "Identification and characterization of a novel MAP kinase kinase  
kinase, MLTK."  
RL J. Biol. Chem. 276:4276-4286(2001).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AB049733; BAB16444.1; -  
DR HSSP; P12931; 1FMK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR001660; SAM.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00536; SAM; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00454; SAM; 1.  
DR SMART; SM00220; STYKc; 1.  
DR SMART; SM00221; STYKc; 1.  
DR SMART; SM00219; TYRKc; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR ATP-binding; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 800 AA; 91188 MW; 2C8593824AB3FADD CRC64;

Query Match 73.7%; Score 1749.5; DB 4; Length 800;  
Best Local Similarity 76.1%; Pred. No. 3.4e-126;  
Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;

QY 1 MSSLGASFVQIKFDDLOFFENC GGSGFSGSVYRAKWI SQDKEVA VKKLKIEKEA EILSVL 60  
Db 1 MSSLGASFVQIKFDDLOFFENC GGSGFSGSVYRAKWI SQDKEVA VKKLKIEKEA EILSVL 60  
QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDYINSNRSEEMDMDHIMTATDVAKGMHY 120  
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDYINSNRSEEMDMDHIMTATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180  
QY 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAVLVVEKNERLTIPSSCPRSFAELLH 240  
Db 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAVLVVEKNERLTIPSSCPRSFAELLH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIEATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIEATLERLKLKLERD 300  
QY 301 LSFKEQELKERERRLKMEQKLTQSNTPLL-LPLAARMSSESYF--ESKTEESNSAE 355  
Db 301 LSFKEQELKERERRLKMEQKLTQSNTPLL-LPLAARMSSESYF--ESKTEESNSAE 355  
QY 356 MSCQITATSNGECHGNPQLQAMLMFGDIFSMN--KAGAVMH--SGMQ-----INM- 404  
Db 356 MSCQITATSNGECHGNPQLQAMLMFGDIFSMN--KAGAVMH--SGMQ-----INM- 404  
QY 405 ---QAKONSSKTSKRGRKGVNMAFGSDFDLSGDDDD-----DGEE 446  
Db 405 ---QAKONSSKTSKRGRKGVNMAFGSDFDLSGDDDD-----DGEE 446  
QY 416 HFPPLIKDSGGEPEENEKIIVLELVFG-FHLKPGTGPDCKWKMYMEMDGE 467  
Db 416 HFPPLIKDSGGEPEENEKIIVLELVFG-FHLKPGTGPDCKWKMYMEMDGE 467

RESULT 5  
Q9HDD2 PRELIMINARY; PRT; 800 AA.  
AC Q9HDD2;  
DT 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, last annotation update)  
DE Plaucible mixed-lineage kinase protein.  
GN MLKLA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymphoid Organ;  
RA Abe Y., Ueda N.;  
RT "Placible Mixed-lineage kinase derived from LAK cell."  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AB030034; BAB12040.1; -  
DR HSSP; P12931; 1FMK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR001660; SAM.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00536; SAM; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00454; SAM; 1.  
DR SMART; SM00221; STYK; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00219; TYRK; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
KW SEQUENCE 800 AA; 9155 MW; B2814509EC54B07A CRC64;  
SQ

Query Match 73.6%; Score 1748.5; DB 4; Length 800;  
Best Local Similarity 76.1%; Pred. No. 4.1e-126;  
Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;

QY 1 MSSLGASFVQIKFDLQFFENCSSGSGSVYRAKWSQDKEVAVKLLKIEKAEILSVL 60  
DB 1 MSSLGASFVQIKFDLQFFENCSSGSGSVYRAKWSQDKEVAVKLLKIEKAEILSVL 60  
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYNINRSEEMDMHMTWATDVAKGMHY 120  
DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYNINRSEEMDMHMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTHTMSLVGTFPMWAEVIOQ 180  
DB 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTHTMSLVGTFPMWAEVIOQ 180  
QY 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240  
DB 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240  
QY 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIEATLERLKLKLERD 300  
DB 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIEATLERLKLKLERD 300  
QY 301 LSFKEQELKERERRLLKMEQKLTQSNTPLL-LPLAARMSEESYF--ESKTEESNSAE 355  
DB 301 LSFKEQELKERERRLLKMEQKLTQSNTPLL-LPLAARMSEESYF--ESKTEESNSAE 355  
QY 356 MSCQITATSNGEHGMNPSLQAMLMGEGDIFSMN--KAGAVMH--SGMQ-----INM- 404  
DB 356 MSCQITATSNGEHGMNPSLQAMLMGEGDIFSMN--KAGAVMH--SGMQ-----INM- 404  
QY 361 MSVYASLFKEN-----NITGRLLLEEDIDKMGIVSKGHIHFKSAIEKLTHTDYINLF 415  
DB 361 MSVYASLFKEN-----NITGRLLLEEDIDKMGIVSKGHIHFKSAIEKLTHTDYINLF 415  
QY 405 ---QAKONSSKTTSKRGGKKNMVALGFSDFDISEGDDDD-----DDGEE 446  
DB 405 ---QAKONSSKTTSKRGGKKNMVALGFSDFDISEGDDDD-----DDGEE 446  
QY 416 HPPPLIKDSGEPENBEKIVNLELVFG-FHLKPGTGPOCDCKWKMYMEMDGE 467  
DB 416 HPPPLIKDSGEPENBEKIVNLELVFG-FHLKPGTGPOCDCKWKMYMEMDGE 467

RESULT 6

Q9NVE9  
ID 09NVE9 PRELIMINARY; PRT; 800 AA.  
AC 09NVE9;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)  
DE 01-JUN-2002 (Tremblrel. 21, last annotation update)  
DE 1 lineage kinase-related kinase MRK-alpha).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Cloning and characterisation of Azk, a mixed lineage kinase  
RA containing a sterile-alpha motif."  
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
RL [2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=21950776; PubMed=11836244;  
RX Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;  
RA "MRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in  
RT gamma-Radiation-Induced Cell Cycle Arrest."  
RL J. Biol. Chem. 277:13873-13882(2002).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AF251441; AAF65822.1; -  
DR EMBL; AF480461; AAL85891.1; -  
DR HSSP; P12931; 1FMK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR001660; SAM.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00536; SAM; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00454; SAM; 1.  
DR SMART; SM00221; STYK; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
KW SEQUENCE 800 AA; 91181 MW; B289D836EC52E295 CRC64;  
SQ

Query Match 73.6%; Score 1748.5; DB 4; Length 800;  
Best Local Similarity 76.1%; Pred. No. 4.1e-126;  
Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;

QY 1 MSSLGASFVQIKFDLQFFENCSSGSGSVYRAKWSQDKEVAVKLLKIEKAEILSVL 60  
DB 1 MSSLGASFVQIKFDLQFFENCSSGSGSVYRAKWSQDKEVAVKLLKIEKAEILSVL 60  
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYNINRSEEMDMHMTWATDVAKGMHY 120  
DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYNINRSEEMDMHMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTHTMSLVGTFPMWAEVIOQ 180  
DB 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTHTMSLVGTFPMWAEVIOQ 180  
QY 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240  
DB 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240  
QY 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIEATLERLKLKLERD 300  
DB 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIEATLERLKLKLERD 300  
QY 301 LSFKEQELKERERRLLKMEQKLTQSNTPLL-LPLAARMSEESYF--ESKTEESNSAE 355  
DB 301 LSFKEQELKERERRLLKMEQKLTQSNTPLL-LPLAARMSEESYF--ESKTEESNSAE 355  
QY 301 LSFKEQELKERERRLLKMEQKLTQSNTPLL-LPLAARMSEESYF--ESKTEESNSAE 355  
DB 301 LSFKEQELKERERRLLKMEQKLTQSNTPLL-LPLAARMSEESYF--ESKTEESNSAE 355

QY 356 :SCQITATSNCEGHGMNPSLOAMLMGFGDIFSMN--KAGAVMH--SGMQ-----INM- 404  
Db 361 MSYVASLFKEN-----NITGRKLLLEEDLKDGMGIVSKGHIHFKSAIEKLTHTDYINLF 415  
QY 405 ---QAKONSSKTSKRGRKKVMALGFSDFDLSEGDGDDDD-----DGEE 446  
Db 416 HFPPLIKDSGGEPEENEKEKIVNLELVFG-FHLKPGTGPDCKMKMYMEMDGE 467

RESULT 7

Q9ESL4  
ID Q9ESL4 PRELIMINARY; PRT; 802 AA.  
AC Q9ESL4;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE MLTK alpha.  
GN ZAK OR MLTK.  
OS Mus musculus (Mouse).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_Taxid=10090;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21264927; PubMed=11042189;  
RA Gotoh I., Adachi M., Nishida E.;  
RT "Identification and Characterization of a Novel MAP Kinase Kinase  
kinase, MLTK.";  
RL J. Biol. Chem. 276:4276-4286(2001).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AB049731; BAB16442.1; .  
DR HSSP; P12931; IFMK.  
DR MGD; MGI:1931274; Zak.  
DR InterPro; IPR000194; ATPase a/bcentre.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR001660; SAM.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00536; SAM; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00454; SAM; 1.  
DR SMART; SM00221; STYKc; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; TYRKc; 1.  
PROSITE; PS00152; ATPASE\_ALPHA\_BETA; UNKNOWN\_1.  
PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
ATP-binding; Serine/threonine-protein kinase; Transferase.  
KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 802 AA; 91719 MW; D431DF8F312A43CC CRC64;

Query Match 73.4%; Score 1743.5; DB 11; Length 802;  
Best Local Similarity 74.3%; Pred. No. 1e-125;  
Matches 356; Conservative 20; Mismatches 58; Indels 45; Gaps 7;

QY 1 MSSLGASFVQIKFDDLOFPENCGGSGFSVYRAKWTSDKEVAVKKLKIEKEAAILSVL 60  
Db 1 MSSLGASFVQIKFDDLOFPENCGGSGFSVYRAKWTSDKEVAVKKLKIEKEAAILSVL 60  
QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMHMTATDVAKGMHY 120  
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMHMTATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIO 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIO 180  
QY 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQYAWLVVEKNERLTIPSSCPRSFAEL 240  
Db 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQYAWLVVEKNERLTIPSSCPRSFAEL 240

QY 241 QCWEADAKKRPSFKOIIISLESMSNDTSLPDKCNSFLHNKAWEKCEIATLERLKL 300  
Db 241 QCWEADAKKRPSFKOIIISLESMSNDTSLPDKCNSFLHNKAWEKCEIATLERLKL 300  
QY 301 LSFKEOELKERERRLKWEOKLTQSNTPPL--LPAAKMSSESYF-----ESKTE 349  
Db 301 LSFKEOELKERERRLKWEOKLTQSNTPPLPSPFETIGAWTEDDVYFWVQOLVRKGE 360  
QY 350 ESNASMSCQITATSNCEGHGMNPSLOAMLMGFGDIFSMN--KAGAVMHSGMQINMOAK 407  
Db 361 MSYVASLFKENNITG-----KRLLEEDLKDGMGIVSKGHIHFKSAIEKLT 409  
QY 408 -----QNSKTSKRGRKKVMALGFSDFDLSEGDGDDDD-----DGEE 446  
Db 410 DYINLHFPPLIKDSGGEPEENEKEKIVNLELVFG-FHLKPGTGPDCKMKMYMEMDGE 467

RESULT 8

Q90ZY8  
ID Q90ZY8 PRELIMINARY; PRT; 371 AA.  
AC Q90ZY8;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Protein kinase Npk.  
OS Brachydanio rerio (Zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
Cyprinidae; Danio.  
NCBI\_Taxid=7955;  
[1]  
RP SEQUENCE FROM N.A.  
RA Chou C.-M., Lee I.-L., Leu J.-H., Huang C.-J.;  
RT "A novel protein kinase, znpk, from the zebrafish."  
Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AF265343; AAK52416.1; .  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
KW ATP-binding; Kinase; Transferase.  
SQ SEQUENCE 371 AA; 42456 MW; 9B918B8A8B20D296 CRC64;

Query Match 59.1%; Score 1402.5; DB 13; Length 371;  
Best Local Similarity 77.6%; Pred. No. 6.4e-100;  
Matches 266; Conservative 32; Mismatches 38; Indels 7; Gaps 2;

QY 1 MSSLGASFVQIKFDDLOFPENCGGSGFSVYRAKWTSDKEVAVKKLKIEKEAAILSVL 60  
Db 32 MSSLGASFVQIKFDDIRFYENC GGSGFSVYRAHWVPQDKEVAVKKLKIDAEAILSVL 91  
QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMHMTATDVAKGMHY 120  
Db 92 SHRNIIOFYGAILEAPNDGIVTEYASRGSLEYESSADSEEMDMQVMTAMEIAKGMHY 151  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIO 180  
Db 152 LHMEAPVKVIHRDLKSRNVVLTADNVLCIDFGASKMVSHTTHMSLVGTFPMAPEVIO 211  
QY 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQYAWLVVEKNERLTIPSSCPRSFAEL 240  
Db 212 LPVSETCDTYSYGVLWEMLTREVPFKGFEGLOQYAWLVVEKHERPTIPSSCPASFADLMR 271  
QY 241 QCWEADAKKRPSFKOIIISLESMSNDTSLPDKCNSFLHNKAWEKCEIATLERLKL 300  
Db 272 RCWNAEPKERPFQKQILGTLETMKNDRLPDQCNFHLNKAWEKCEIETLERLKL 331  
QY 301 LSFKEOELKERERRLKWEOKLTQSN--TPLLPLAARMSE 341  
Db 332 LSCKEQELERERRLTEWENRLMERSRGCTPVSL-----ASEE 369





ID Q9H1Y7 PRELIMINARY; PRT; 564 AA.  
AC Q9H1Y7;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE DJ862P8.3 (Similar to MAP3K10 (Mitogen-activated protein kinase kinase kinase 10)) (Fragment).  
GN DJ862P8.3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hall R.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
DR EMBL; AL133380; CAC17571.1; -.  
DR HSSP; P06241; 1SHF.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS50011; PROTEIN KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50002; SH3; 1.  
KW Kinase; SH3 domain.  
FT NON\_TER  
SQ SEQUENCE 564 AA; 62400 MW; D4F91E14842E7EEB CRC64;

Query Match 25.2%; Score 597.5; DB 4; Length 564;  
Best Local Similarity 37.4%; Pred. No. 1.1e-37;  
Matches 135; Conservative 66; Mismatches 115; Indels 45; Gaps 7;

QY 6 ASFVQIKFDDLOFFENCSSGSGSVYRAKVISQDKEVAVKLLK-----IEKE 53  
DB 114 SSPVHVAFERLEKELIGAGFGQVYRATW--QGQEVAVKARQDPEQDAAAASVRR 171  
QY 54 AEILSVLSHRNIIQFYGVILEPPNYGIYTEYASLSGLYDI-----NSNRSEE 101  
DB 172 ARLFAMLRHPNIIELRGVCLQOPHLCLVLEFARGGALNRALAAANAPDPRAPRRARR 231  
QY 102 MDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVI-----AADGVLCIDFG 153  
DB 232 IPPHVLVNWAVQIARGMVLYHEEAFVPIILHRDLKSSNILLLEKIEHDDICNKTLLKITDFG 291  
QY 154 ASRFHNHTHMSLVGTFPMAPEVIOSLPVSETCDTYSYGVVLMWMLTREVPFGLEGLO 213  
DB 292 LAREWHRTTKMSTAGTYAMWAPAEVIKSLFSKSDIWSYGVLLWELLTGEVPRGIDGLA 351  
QY 214 VAWLVEKNERLTIPSSCPRSFAELLHQWEADAKKRPSPFOIISILESMSND--TSLPD 271  
DB 352 VAYGVAVNKLTLPISCTCEPFAKLMEKCWQODPHIRPSFALLIEQLTAIEGAVMTEMPQ 411  
QY 272 KNSFLHNKAWEKCEIATLERLKLRLDSFKEQL-----KERERLKMWEOKLTE 324  
DB 412 E--SFHSMQDDWKLIEIQMFDELRTKEKELRSREELTRALQOKSQEELLKRRREQOLAE 469  
QY 325 Q 325  
DB 470 R 470

RESULT 12  
ID Q8WWN1 PRELIMINARY; PRT; 1036 AA.  
AC Q8WWN1;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Mixed lineage kinase 4beta.  
GN MLK4BETA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kvashe S., Protodopov A., Rynditch A., Zabarovsky E., Kashuba V.,  
RT "MLK4, a new member of mixed lineage kinases."  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ311798; CAC84640.1; -.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN\_1.  
DR PROSITE; PS50011; PROTEIN KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
DR PROSITE; PS50002; SH3; 1.  
KW Kinase.  
SQ SEQUENCE 1036 AA; 113809 MW; B9C2ACF397C2CF37 CRC64;

Query Match 24.9%; Score 590.5; DB 4; Length 1036;  
Best Local Similarity 33.1%; Pred. No. 8.7e-37;  
Matches 149; Conservative 75; Mismatches 147; Indels 79; Gaps 9;

QY 6 ASFVQIKFDDLOFFENCSSGSGSVYRAKVISQDKEVAVKLLK-----IEKE 53  
DB 114 SSPVHVAFERLEKELIGAGFGQVYRATW--QGQEVAVKARQDPEQDAAAASVRR 171  
QY 54 AEILSVLSHRNIIQFYGVILEPPNYGIYTEYASLSGLYDI-----NSNRSEE 101  
DB 172 ARLFAMLRHPNIIELRGVCLQOPHLCLVLEFARGGALNRALAAANAPDPRAPRRARR 231  
QY 102 MDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVI-----AADGVLCIDFG 153  
DB 232 IPPHVLVNWAVQIARGMVLYHEEAFVPIILHRDLKSSNILLLEKIEHDDICNKTLLKITDFG 291  
QY 154 ASRFHNHTHMSLVGTFPMAPEVIOSLPVSETCDTYSYGVVLMWMLTREVPFGLEGLO 213  
DB 292 LAREWHRTTKMSTAGTYAMWAPAEVIKSLFSKSDIWSYGVLLWELLTGEVPRGIDGLA 351  
QY 214 VAWLVEKNERLTIPSSCPRSFAELLHQWEADAKKRPSPFOIISILESMSND--TSLPD 271  
DB 352 VAYGVAVNKLTLPISCTCEPFAKLMEKCWQODPHIRPSFALLIEQLTAIEGAVMTEMPQ 411  
QY 272 KNSFLHNKAWEKCEIATLERLKLRLDSFKEQL-----KERERLKMWEOKLTE 324  
DB 412 E--SFHSMQDDWKLIEIQMFDELRTKEKELRSREELTRALQOKSQEELLKRRREQOLAE 469  
QY 325 Q-----SNTPLLLPLAARMSSESYFESKTESNSA 354  
DB 470 REIDVLERELNILLIQLNQEKPVKYKRGKFKRSRLKLDGHRISLPSDFQHKITVQASP 529  
QY 355 EM-----SCOITATSNGEHGMNPSLOAMML 380  
DB 530 NLDKRRSLNSSSSSPSSPTMPLRAIQL 559

[illegible]

```

RESULT 14
ID Q16584 PRELIMINARY; PRT; 847 AA.
AC Q16584;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE protein kinase (Similar to mitogen-activated protein kinase kinase
de kinase 11).
GN MLK-3 OR SPRK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=THYMUS;
RX MEDLINE=94239754; PubMed=8183572;
RA Ing Y.L., Leung I.W., Heng H.H., Tsui L.C., Lassam N.J.;
RT "MLK-3: identification of a widely-expressed protein kinase bearing an
RT SH3 domain and a leucine zipper-basic region domain.";
RL Oncogene 9:1745-1750(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94253068; PubMed=8195146;
RA Gallo K.A., Mark M.R., Scadden D.T., Wang Z., Gur Q., Godowski P.J.;
RT "Identification and characterization of SPRK, a novel src-homology 3
RT domain-containing proline-rich kinase with serine/threonine kinase
RT activity.";
RL J. Biol. Chem. 269:15092-15100(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; L32976; AAA59859.1; -.
DR EMBL; U07747; AAA19647.1; -.
DR EMBL; BC011263; AAH11263.1; -.
DR HSSP; P29355; ISEM.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TYRKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50002; SH3; 1.
KW ATP-binding; Kinase; SH3 domain; Serine/threonine-protein kinase;
KW Transferase.
KW SEQUENCE 847 AA; 92687 MW; AFB6E930EA281C15 CRC64;

Query Match 24.5%; Score 581.5; DB 4; length 847;
Best Local Similarity 37.4%; Pred. No. 3.3e-36;
Matches 132; Conservative 60; Mismatches 122; Indels 39; Gaps

QY 13 FDDLQFFENGCGGSGFVSYPRAKWISQ-----DKEVAVKLLIKIEKAELSVLS 61
DB 114 FQELRLAEVIGTIGRGKVVYRGSWMRGELVAVKARQDPDEDISV-TAESVRQEARLFAMLA 172

QY 62 HRNIIQFYGVILPEPNYGVITEYASLSGLYDINSNRSEEMDMDHIM-TWATDVAKGMHY 120
DB 173 HPNIIALKAVCLEENLCLVMEYAAGGPLRALAGRVP---PHVLVNWAVQIARGMHY 228

```

QY 121 LHMEAPVYVIHRDLKSRNV-----IAADGV-----LKICDFGASRFHNHTTHMSLVGTFFW 172  
Db 229 LHCEALVPIVHRDLKSNLILQPIESDDMEHKTLDGLAREWHKTQMSAAGTYAW 288  
QY 173 MAPEVIQSLPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVLVEKNERLTISSCP 232  
Db 289 MAPEVIKASTFSKGSVDWSPFGLVLELLTGEVPRGIDCLAVAYGAVNKLTLPISTCP 348  
QY 233 RSFAELLHQCWEADAKKRPSPFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLE 292  
Db 349 EPFAQLMADCWADPHRRPDPFASILQQLLEALQVLRMPRDSFHSMQEGWKREIQGLFD 408  
QY 293 RLKKLERDLSFKEOE-----LKERERRLKMWEOKLTEQSNTPPL 331  
Db 409 ELRAKEKELLSREBELTRAAREQRSQAEOQLRRREHLLAQWELVFERDVTLL 461

RESULT 15

Q9J15 PRELIMINARY; PRT; 850 AA.  
AC Q9J15; 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Mixed lineage kinase 3.  
GN MLK3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/OLA;  
RX MEDLINE=20354997; Pubmed=10894943;  
RA Saridaki A., Ferraz C., Demaille J., Scherer G., Roux A.-F.;  
RT "Genomic sequencing reveals the structure of the Kcnk6 and Map3k11  
RT genes and their close vicinity to the Sipal gene on mouse chromosome  
RT 19.";  
RL Cytogenet. Cell Genet. 89:85-88(2000).  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
DR EMBL; AF155142; AAF73281.1; -.  
DR HSSP; P29355; 1SEM.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00219; TYRK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50002; SH3; 1.  
KW ATP-binding; Kinase; SH3 domain; Serine/threonine-protein kinase;  
KW Transferase.  
SO SEQUENCE 850 AA; 93199 MW; 8F026CB3532DC10E CRC64;

Query Match 24.4%; Score 580.5; DB 11; Length 850;  
Best Local Similarity 37.4%; Pred. No. 3.9e-36;  
Matches 102; Conservative 59; Mismatches 123; Indels 39; Gaps 7;

QY 13 FDDLQFFENGCGGSGFSGVYRAKWSQ-----DKEVAVKLLKIEKAELSVLS 61  
Db 115 FQELRLLEEVIGIGFGKYYRGSWRGELVAVKARQDDEDISV-TAESVQEARLFLAMLA 173  
QY 62 HRNIIOFYGVILEPPNYGIVTEYASLGLDYDINSNRSEMDMDHM-TWATDVAKGMHY 120

Db 174 HPNIIALKAIVCLEEPNLCVMEYAAAGPLSRALAGRVP-----PHVLVNWAVQIARGMHY 229  
QY 121 LHMEAPVYVIHRDLKSRNV-----IAADGV-----LKICDFGASRFHNHTTHMSLVGTFFW 172  
Db 230 LHCEALVPIVHRDLKSNLILQPIEGDDMEHKTLDGLAREWHKTQMSAAGTYAW 289  
QY 173 MAPEVIQSLPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVLVEKNERLTISSCP 232  
Db 290 MAPEVIKASTFSKGSVDWSPFGLVLELLTGEVPRGIDCLAVAYGAVNKLTLPISTCP 349  
QY 233 RSFAELLHQCWEADAKKRPSPFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLE 292  
Db 350 EPFAQLMADCWADPHRRPDPFASILQQLLEALQVLRMPRDSFHSMQEGWKREIQGLFD 409  
QY 293 RLKKLERDLSFKEOE-----LKERERRLKMWEOKLTEQSNTPPL 331  
Db 410 ELRAKEKELLSREBELTRAAREQRSQAEOQLRRREHLLAQWELVFERDVTLL 462

Search completed: May 1, 2003, 20:37:45  
Job time : 58.7671 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:12:12 ; Search time 87.5732 Seconds  
(without alignments)  
692.324 Million cell updates/sec

Title: US-09-757-982-5  
Perfect score: 2375  
Sequence: 1 MSSLGASFVQIKFDDLOFFE.....GDDDDDDGEEEDNDMDNSE 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description                     |
|------------|--------|-------------|--------|-------|---------------------------------|
| 1          | 2375   | 100.0       | 455    | 21    | AAB18657 A human regulator      |
| 2          | 2375   | 100.0       | 455    | 21    | AAV83278 Human survival reg     |
| 3          | 2375   | 100.0       | 455    | 21    | AAV84321 A human cardiovascular |
| 4          | 2366   | 99.6        | 473    | 22    | AAW25322 Human protein sequ     |
| 5          | 1748.5 | 73.6        | 800    | 22    | AAB71957 Human TGF-beta rec     |
| 6          | 1748.5 | 73.6        | 800    | 22    | AAB65673 Novel protein kina     |
| 7          | 1409   | 59.3        | 349    | 22    | AAV75571 Human colon cancer     |
| 8          | 744    | 31.3        | 141    | 21    | AAV03583 Human secreted pro     |
| 9          | 638    | 26.9        | 124    | 22    | AAB65552 C-terminus specifi     |
| 10         | 604    | 25.4        | 1021   | 23    | ABP61000 Novel human protei     |

|    |       |      |      |    |                             |
|----|-------|------|------|----|-----------------------------|
| 11 | 599.5 | 25.2 | 719  | 22 | AAB85513 Human protein kina |
| 12 | 599.5 | 25.2 | 1036 | 23 | ABB80923 Novel human protei |
| 13 | 590   | 24.8 | 1097 | 23 | AAE21717 Human PKIN-12 prot |
| 14 | 589   | 24.8 | 1046 | 22 | AAE11775 Human kinase (PKIN |
| 15 | 581.5 | 24.5 | 847  | 23 | AAE22763 Human leucine acti |
| 16 | 565.5 | 23.8 | 859  | 16 | AAR82886 Human leucine zipp |
| 17 | 565.5 | 23.8 | 859  | 18 | AAW31227 Human leucine-zipp |
| 18 | 564.5 | 23.8 | 888  | 23 | ABB57049 Mouse ischaemic co |
| 19 | 544.5 | 22.9 | 1020 | 22 | ABB58999 Drosophila melanog |
| 20 | 527.5 | 22.2 | 977  | 22 | ABB71694 Novel human diagno |
| 21 | 510.5 | 21.5 | 1490 | 22 | ABG19123 Melon constitutive |
| 22 | 504.5 | 21.2 | 850  | 23 | AAE18529 Tomato TCTR2. Lyc  |
| 23 | 495   | 20.8 | 982  | 22 | AAB50439 Rice EDRL. Oryza   |
| 24 | 494   | 20.8 | 903  | 22 | AAB50440 Arabidopsis CTPI p |
| 25 | 488   | 20.5 | 821  | 16 | AAR80574 Constitutive tripl |
| 26 | 488   | 20.5 | 821  | 18 | AAW17938 Arabidopsis thalia |
| 27 | 488   | 20.5 | 821  | 22 | AAB50438 Arabidopsis thalia |
| 28 | 484   | 20.4 | 821  | 15 | AAR46723 Mouse transforming |
| 29 | 483   | 20.3 | 579  | 18 | AAW27092 Human TGF-beta act |
| 30 | 483   | 20.3 | 579  | 20 | AAV28996 Human TGF-beta act |
| 31 | 483   | 20.3 | 579  | 23 | ABB85033 Human TAK1 protein |
| 32 | 481   | 20.3 | 579  | 18 | AAW27093 Human TAK-1 protei |
| 33 | 481   | 20.3 | 579  | 20 | AAV09542 Human TAK1-6XHis p |
| 34 | 481   | 20.3 | 579  | 21 | AAV91000 Arabidopsis thalia |
| 35 | 481   | 20.3 | 590  | 20 | AAV09547 Arabidopsis thalia |
| 36 | 478.5 | 20.1 | 933  | 22 | AAB50437 Arabidopsis thalia |
| 37 | 478.5 | 20.1 | 970  | 22 | AAB50443 Arabidopsis thalia |
| 38 | 475.5 | 20.0 | 369  | 21 | AAV22172 Human TGF-beta act |
| 39 | 475.5 | 20.0 | 374  | 21 | AAV22171 Human TGF-beta act |
| 40 | 475.5 | 20.0 | 412  | 21 | AAV22170 Human TGF-beta act |
| 41 | 471.5 | 19.9 | 567  | 20 | AAV28998 Human TGF-beta act |
| 42 | 471.5 | 19.9 | 606  | 20 | AAV28997 Novel human diagno |
| 43 | 470   | 19.8 | 144  | 22 | ABG06092 Drosophila melanog |
| 44 | 468.5 | 19.7 | 678  | 22 | ABB58061 Arabidopsis thalia |
| 45 | 447.5 | 18.8 | 367  | 21 | AAV32053                    |

ALIGNMENTS

|  |   |
|--|---|
| RESULT 1   |   |
| AAB18657   |   |
| ID AAB18657 standard; Protein; 455 AA.                                   |   |
| XX AAB18657;   |   |
| AC   |   |
| XX   |   |
| DT 22-JAN-2001 (first entry)   |   |
| XX   |   |
| DE A human regulator of intracellular phosphorylation.                   |   |
| XX   |   |
| KW Human; intracellular phosphorylation regulator; HRP; stroke; myeloma; |   |
| KW neurological disorder; Parkinson's disease; demyelinating disease;    |   |
| KW meningitis; developmental disorder; neuromuscular disorder; cancer;   |   |
| KW myasthenia gravis; cell proliferative disorder; actinic keratosis;    |   |
| KW arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;   |   |
| KW autoimmune disorder; inflammatory disorder; Addison's disease;        |   |
| KW acquired immunodeficiency disease; allergy; diabetes mellitus;        |   |
| KW rheumatoid arthritis; microbial infection; trauma.                    |   |
| XX   |   |
| OS Homo sapiens.   |   |
| XX   |   |
| FH   |   |
| FT Key   | Location/Qualifiers                       |
| FT Domain  | 16...257                                  |
| FT   | /note= "eukaryotic protein kinase domain" |
| FT   | 61  |
| FT   | /note= "potential phosphorylation site"   |
| FT   | 89  |
| FT   | /note= "potential phosphorylation site"   |
| FT   | 96  |
| FT   | /note= "potential phosphorylation site"   |
| FT   | 97  |
| FT   | /note= "potential glycosylation site"     |
| FT   | 129..141                                  |
| FT   | Binding-site                              |

FT Modified-site /note= "protein kinase ATP-binding site"  
FT 159 /note= "potential glycosylation site"  
FT 234 /note= "potential phosphorylation site"  
FT 252 /note= "potential phosphorylation site"  
FT 258 /note= "potential phosphorylation site"  
FT /note= "potential phosphorylation site"  
FT 265 /note= "potential glycosylation site"  
FT 268 /note= "potential phosphorylation site"  
FT /note= "potential phosphorylation site"  
FT 294..322 /note= "leucine zipper"  
FT 302 /note= "potential phosphorylation site"  
FT /note= "potential phosphorylation site"  
FT 302 /note= "potential phosphorylation site"  
FT /note= "potential phosphorylation site"  
FT 342 /note= "potential phosphorylation site"  
FT /note= "potential phosphorylation site"  
FT 343 /note= "potential phosphorylation site"  
FT /note= "potential phosphorylation site"  
FT 346 /note= "potential phosphorylation site"  
FT /note= "potential phosphorylation site"  
FT 364 /note= "potential phosphorylation site"  
FT 409 /note= "potential glycosylation site"  
FT /note= "potential glycosylation site"  
FT 410 /note= "potential phosphorylation site"  
FT /note= "potential phosphorylation site"  
FT 414 /note= "potential phosphorylation site"  
FT /note= "potential phosphorylation site"  
FT 415 /note= "potential phosphorylation site"  
FT /note= "potential phosphorylation site"  
FT 429 /note= "potential phosphorylation site"  
FT /note= "potential phosphorylation site"  
FT 434 /note= "potential phosphorylation site"  
FT /note= "potential phosphorylation site"  
FT WO200055332-A2.  
XX  
XX  
PD 21-SEP-2000.  
XX  
PF 17-MAR-2000; 2000WO-US07277.  
XX  
PR 18-MAR-1999; 99US-0125593.  
20-MAY-1999; 99US-0135049.  
09-JUL-1999; 99US-0143188.  
PA (INCY-) INCYTE PHARM INC.  
XX Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;  
PI Lu DAM, Au-Young J;  
XX  
XX WPI; 2000-602121/57.  
DR N-PSDB; AAA75674.  
XX  
XX Novel human intracellular phosphorylation regulator polypeptides and  
PT polynucleotides for diagnosis, prevention and treatment of  
PT neurological, cell proliferative and autoimmune/inflammatory disorders  
PT  
XX  
PS Claim 1; Page 75-76; 96pp; English.  
XX  
XX The present sequence represents a human regulator of intracellular  
CC phosphorylation (HRIP). HRIP is useful for screening agonists and  
CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist  
CC are useful for treating a disease or condition associated with  
CC decreased or increased expression of functional HRIP. Diseases treated  
CC or diagnosed include neurological disorders such as stroke, Parkinson's  
CC disease, demyelinating diseases, bacterial and viral meningitis and  
CC other developmental disorders of the central nervous system,  
PA

CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders  
CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer  
CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,  
CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/  
CC inflammatory disorder such as Addison's disease, acquired  
CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,  
CC rheumatoid arthritis, microbial infection and trauma.  
XX  
SQ Sequence 455 AA;  
Query Match 100.0%; Score 2375; DB 21; length 455;  
Best Local Similarity 100.0%; Pred. No. 6.5e-207;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSLGASFVQIKEDDLQFFENCGGSGFSGYRAKWISQDKAVAKLLIKKBAEILSVL 60  
Db 1 MSLGASFVQIKEDDLQFFENCGGSGFSGYRAKWISQDKAVAKLLIKKBAEILSVL 60  
QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLGLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLGLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFPMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFPMAPEVIQS 180  
QY 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPSPFAELH 240  
Db 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPSPFAELH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCIEATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCIEATLERLKLKLERD 300  
QY 301 LSPKEQELKERERRLKWMEQKLTQOSNTPLILPLARMSEESYFESKTEESNSAEMSCOI 360  
Db 301 LSPKEQELKERERRLKWMEQKLTQOSNTPLILPLARMSEESYFESKTEESNSAEMSCOI 360  
QY 361 TATSNGEHGMPNPSLOAMLMGFGDIFSMNKAAGVMSGMQINMOAKONSSKTSKRGRK 420  
Db 361 TATSNGEHGMPNPSLOAMLMGFGDIFSMNKAAGVMSGMQINMOAKONSSKTSKRGRK 420  
QY 421 KVNMAIGFSDFDLSEGD DDDDDGEEEDNDMDNSE 455  
Db 421 KVNMAIGFSDFDLSEGD DDDDDGEEEDNDMDNSE 455  
RESULT 2  
ID AAY83278 standard; Protein; 455 AA.  
XX AAY83278;  
AC  
XX  
DT 16-AUG-2000 (first entry)  
XX  
XX Human survival regulating kinase (SRK).  
XX  
XX Survival regulating kinase; SRK; Raf; MBP; BAD; MAPKK; MAPKK;  
KW HAX-1; protein kinase; autophosphorylation; cell growth; regulation;  
KW apoptosis; cell survival; nuclear targeting; tumour; human;  
KW autoimmune disease.  
OS Homo sapiens.  
XX  
XX WO200022142-A2.  
PN  
XX  
PD 20-APR-2000.  
XX  
XX 20-SEP-1999; 99WO-US22008.  
PF  
XX  
XX 13-OCT-1998; 98US-0104088.  
PR  
XX  
XX (ONYX-) ONYX PHARM INC.  
PA

XX Ruggieri R, Callow M, Diaz P,  
XX WPI, 2000-317994/27.  
DR N-PSDB; AA293783.  
XX  
PT Novel human survival regulating kinase polypeptide for screening agents  
PT which modulate biological pathways associated with SRK useful in  
PT treating autoimmune diseases, tumors and apoptosis-related disorders  
PS Claim 4; Figure 2; 62pp; English.  
XX  
CC Survival regulating kinases (SRK) are a class of proteins involved in  
CC cell signal transduction pathways such as mitogen-activated protein  
CC kinase pathways. A protein kinase activity means that the SRK can  
CC catalyze a reaction in which a phosphate group is transferred from a  
CC phosphate donor to a phosphate acceptor amino acid residue,  
CC preferably the hydroxyl side chain of a serine or threonine.  
CC Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase  
CC activity is similar to that of a MAPKK such as Raf. has a range of  
CC other activities including a cell growth-regulatory activity, a cell  
CC survival promoting activity, a HAX-1 binding activity, an apoptosis  
CC suppressing activity a MAPK activation or stimulatory activity, a  
CC nuclear targeting activity and a SRK-specific immunogenic activity.  
CC SRK is useful for identifying agents which modulate cellular  
CC transformations mediated by Ras and SRK and agents that modulate the  
CC apoptosis suppression activity of SRK. This information may be useful  
CC in the treatment of autoimmune diseases, tumors and apoptosis  
CC related disorders.  
XX  
SQ Sequence 455 AA;  
Query Match 100.0%; Score 2375; DB 21; Length 455;  
Best Local Similarity 100.0%; Pred. No. 6.5e-207;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSQDKEAVAVKKLKEKEAEILSVL 60  
Db 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSQDKEAVAVKKLKEKEAEILSVL 60  
QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180  
181 LPVSETCDTYSYGVLWEMLTREVVPFKGLEGLQVAVLVEKNERLTIPSSCPRSPFAELH 240  
Db 181 LPVSETCDTYSYGVLWEMLTREVVPFKGLEGLQVAVLVEKNERLTIPSSCPRSPFAELH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIATLERLKLKLERD 300  
QY 301 LSFKEQELKERERRLKWEQKLTQSNTPLLPLAARMESESYFESKTESNSAEMSCQI 360  
Db 301 LSFKEQELKERERRLKWEQKLTQSNTPLLPLAARMESESYFESKTESNSAEMSCQI 360  
QY 361 TATSNGEGHGMNPSLQAMLMGFGDIFSMMKAGAVHSGMOINMOAKONSSKTTSKRGK 420  
Db 361 TATSNGEGHGMNPSLQAMLMGFGDIFSMMKAGAVHSGMOINMOAKONSSKTTSKRGK 420  
QY 421 KVNMALGFSDFDLSEGD DDDDDGEEEDNDMDNSE 455  
Db 421 KVNMALGFSDFDLSEGD DDDDDGEEEDNDMDNSE 455  
RESULT 3  
ID AAY84321 standard; Protein; 455 AA.  
XX

AC AAY84321;  
XX  
DT 12-JUL-2000 (first entry)  
XX  
DE A human cardiovascular system associated protein kinase-2.  
XX  
KW Human; cardiovascular system associated protein kinase-2; CSAPK-2;  
KW signaling pathway; cell growth; cell differentiation; gene mapping;  
KW tissue typing; forensic identification; cardiovascular disease;  
KW congestive heart failure; transgenic animal.  
XX  
OS Homo sapiens.  
XX  
PN WO200014212-A1.  
XX  
PD 16-MAR-2000.  
XX  
PF 09-SEP-1999; 99WO-US20631.  
XX  
PR 09-SEP-1998; 98US-0099657.  
PR 29-SEP-1998; 98US-0163115.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Acton S;  
XX  
DR WPI, 2000-271053/23.  
DR N-PSDB; AA299726, AA299727.  
XX  
PT New nucleic acid encoding cardiovascular system associated protein  
PT kinase, used e.g. for diagnosis, treatment and prevention of  
PT cardiovascular disease  
XX  
PS Claim 2; Fig 2; 163pp; English.  
XX  
CC The present sequence represents a human cardiovascular system associated  
CC protein kinase-2 (CSAPK-2). CSAPK polypeptides are involved in signalling  
CC pathways associated with cell growth and differentiation. The CSAPK  
CC polypeptides and polynucleotides are used to screen for agents that  
CC specifically modulate CSAPK, which are potential therapeutic agents.  
CC They are also used for diagnosis, prognosis or monitoring of  
CC CSAPK-related diseases, gene mapping, tissue typing and forensic  
CC identification, and for treating or preventing disorders associated  
CC with aberrant CSAPK expression or activity, especially cardiovascular  
CC diseases such as congestive heart failure. They can also be used in  
CC pharmacogenomics. The CSAPK polynucleotide may also be used to generate  
CC transgenic animals.  
XX  
SQ Sequence 455 AA;  
Query Match 100.0%; Score 2375; DB 21; Length 455;  
Best Local Similarity 100.0%; Pred. No. 6.5e-207;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSQDKEAVAVKKLKEKEAEILSVL 60  
Db 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSQDKEAVAVKKLKEKEAEILSVL 60  
QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180  
QY 181 LPVSETCDTYSYGVLWEMLTREVVPFKGLEGLQVAVLVEKNERLTIPSSCPRSPFAELH 240  
Db 181 LPVSETCDTYSYGVLWEMLTREVVPFKGLEGLQVAVLVEKNERLTIPSSCPRSPFAELH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIATLERLKLKLERD 300

QY 301 LSFKEQLKERERRLKWEOKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQI 360  
DB 301 LSFKEQLKERERRLKWEOKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQI 360  
QY 361 TATSNGEHGMNPSLOAMMLMGFGDIFSMNKAGAVMHSQMGINMOAKONSSKTSKRGRK 420  
DB 361 TATSNGEHGMNPSLOAMMLMGFGDIFSMNKAGAVMHSQMGINMOAKONSSKTSKRGRK 420  
QY 421 KVMNALGFSDFDLSEGDGDDDDDDGEEEDNDMDNSE 455  
DB 421 KVMNALGFSDFDLSEGDGDDDDDDGEEEDNDMDNSE 455  
RESULT 4  
ID AAM25322  
AC AAM25322 standard; Protein; 473 AA.  
XX 16-OCT-2001 (first entry)  
DE Human protein sequence SEQ ID NO:837.  
XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KW antinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
KW antibacterial; endocrine; cardiant; central nervous system; virucide;  
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
KW antiaggregant; haemostatic; vulnery; antiulcer; osteopathic; eczema;  
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;  
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder.  
XX Homo sapiens.  
OS WO200153455-A2.  
XX 26-JUL-2001.  
PD 22-DEC-2000; 2000WO-US35017.  
PF 23-DEC-1999; 99US-0471275.  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
XX (HYSE-).HYSEQ INC.  
XX PA Tang YT, Liu C, Drmanac RT;  
XX PI WPI; 2001-457603/49.  
XX DR N-PSDB; AAH99263.  
XX DR  
XX PT Isolated human polynucleotides encoding polypeptides, useful for the  
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -  
XX  
XX Claim 20; Page 191; 1217pp; English.  
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
CC AAM25963. The proteins can have activities based on the tissues and  
CC cells they are expressed in, such as: antinflammatory; antirheumatic;  
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;  
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;  
CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;  
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;  
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides

CC encoding them can be used in gene therapy, antisense therapy and vaccine  
CC production. The proteins and polynucleotides are useful for screening for  
CC agonists or antagonists of a protein and for the treatment e.g. inflammation,  
CC of disorders associated with the activity of a protein e.g. inflammation,  
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders.  
XX  
SQ Sequence 473 AA;  
Query Match 99.6%; Score 2366; DB 22; Length 473;  
Best Local Similarity 99.8%; Pred. No. 4.6e-206;  
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MSSLGASFVQIKFDDLOFFENCGGSGSVYRAKWTISODKEVAVKKLTKIEKAEILSVL 60  
DB 19 MSSLGASFVQIKFDDLOFFENCGGSGSVYRAKWTISODKEVAVKKLTKIEKAEILSVL 78  
QY 61 SHRNIQFYGVILEPPNYGIVTEYASLSGLYDYNNSRSEMDMDHIMTWATVAKGMHY 120  
DB 79 SHRNIQFYGVILEPPNYGIVTEYASLSGLYDYNNSRSEMDMDHIMTWATVAKGMHY 138  
QY 121 LHMEAPVKYIHRDLKSRNVVIADGVLIKIDFGASRFHNHTTHMSLVGTFPMWAPVIO 180  
DB 139 LHMEAPVKYIHRDLKSRNVVIADGVLIKIDFGASRFHNHTTHMSLVGTFPMWAPVIO 198  
QY 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAVLVVEKNERTLTPSSCPSPFAELH 240  
DB 199 LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAVLVVEKNERTLTPSSCPSPFAELH 258  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEWRCEIATLERLKLERD 300  
DB 259 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEWRCEIATLERLKLERD 318  
QY 301 LSFKEQLKERERRLKWEOKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQI 360  
DB 319 LSFKEQLKERERRLKWEOKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQI 378  
QY 361 TATSNGEHGMNPSLOAMMLMGFGDIFSMNKAGAVMHSQMGINMOAKONSSKTSKRGRK 420  
DB 379 TATSNGEHGMNPSLOAMMLMGFGDIFSMNKAGAVMHSQMGINMOAKONSSKTSKRGRK 438  
QY 421 KVMNALGFSDFDLSEGDGDDDDDDGEEEDNDMDNSE 455  
DB 439 KVMNALGFSDFDLSEGDGDDDDDDGEEEDNDMDNSE 473  
RESULT 5  
ID AAB71957  
AC AAB71957 standard; Protein; 800 AA.  
XX AAB71957;  
XX 11-MAY-2001 (first entry)  
XX Human TGF-beta receptor encoded by cDNA clone HDPSM48.  
XX  
XX Human; antisclerotic; dermatological; immunosuppressive; cytostatic;  
KW antinflammatory; anti-HIV; immunostimulant; cardiant; vascular;  
KW ophthalmological; neuroprotective; nootropic; anticonvulsant; vaccine;  
KW antiparkinsonian; antimicrobial; vulnery; gene therapy; infection;  
KW transforming growth factor; TGF; TGF-beta receptor; immune disorder;  
KW hyperproliferative disorder; cardiovascular disease; angiogenesis;  
KW neurological disorder.  
XX Homo sapiens.  
OS WO200112670-A1.  
XX PN



XX 22-FEB-2001.  
PD  
XX  
XX 10-AUG-2000; 2000WO-US21736.  
PF  
XX 13-AUG-1999; 99US-0148682.  
PR 20-SEP-1999; 99US-0154887.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Ni J;  
XX  
DR WPI; 2001-202858/20.  
DR N-PSDB; AAF75336.  
XX  
PT Nucleic acid molecules encoding 12 transforming growth factor-beta  
PT receptor polypeptides, useful for preventing, diagnosing and treating  
PT e.g. cancers, Parkinson's disease and diabetic retinopathy -  
XX

Claim 11; Page 293-295; 311pp; English.

CC The present sequence is one of 12 novel human transforming growth factor  
CC (TGF)-beta receptor polypeptides. The TGF-beta receptor polynucleotides  
CC and polypeptides may be used in the prevention, diagnosis and treatment  
CC of diseases associated with inappropriate polypeptide expression. Such  
CC diseases include immune disorders (e.g. multiple sclerosis, systemic  
CC lupus erythematosus and human immuno-deficiency virus (HIV) infections),  
CC hyperproliferative disorders (e.g. cancers and Gaucher's disease),  
CC cardiovascular diseases (e.g. Scimitar syndrome, Chaga's cardiomyopathy  
CC and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft  
CC neovascularisation and diabetic retinopathy), neurological disorders  
CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease)  
CC and infectious diseases. The polynucleotides and polypeptides are also  
CC useful for promoting wound healing, regeneration and/or chemotaxis. The  
CC polynucleotides and their complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar nucleic acid sequences in samples. The polypeptides may be used  
CC as antigens in the production of antibodies and in assays to identify  
CC modulators of protein expression and activity. The anti-TGF-beta receptor  
CC antibodies may be used to down regulate expression and activity and as  
CC diagnostic agents for detecting the presence of the polypeptides in  
CC samples.  
XX  
XX

Sequence 800 AA;

Query Match

Best Local Similarity 73.6%; Score 1748.5; DB 22; Length 800;

Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;

1 MSSLGASFVQIKFDDLOFFENCSSGFSVYRAKMWISQDKEVAVKLLKIEKEATILSVL 60  
1 MSSLGASFVQIKFDDLOFFENCSSGFSVYRAKMWISQDKEVAVKLLKIEKEATILSVL 60  
Db 1 MSSLGASFVQIKFDDLOFFENCSSGFSVYRAKMWISQDKEVAVKLLKIEKEATILSVL 60  
QY 61 SHRNIQFYGVILEPNNYGIYEVASLSGLYDINSNRSEEMDMDHIMTWATVAKGMHY 120  
Db 61 SHRNIQFYGVILEPNNYGIYEVASLSGLYDINSNRSEEMDMDHIMTWATVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMWAEVIOQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMWAEVIOQS 180  
QY 181 LPVSETCDTYSYGVNLWEMLTREVFPKLEGLQVAVLVEKNERLTIPSSCPSPFAELLH 240  
Db 181 LPVSETCDTYSYGVNLWEMLTREVFPKLEGLQVAVLVEKNERLTIPSSCPSPFAELLH 240  
QY 241 QCWEADAKRPSFKQIISLESMSNDTSLPDKCNFPHNKAEMRCEIATLERLKLIERD 300  
Db 241 QCWEADAKRPSFKQIISLESMSNDTSLPDKCNFPHNKAEMRCEIATLERLKLIERD 300  
QY 301 LSFKEQLKERERRRLKMEQKLTQOSNTPLL-LPLAARMSEESYF--ESKTEESNSAE 355  
Db 301 LSFKEQLKERERRRLKMEQKLTQOSNTPLLPSFEIGAWTEDDVYCWVQQLVRAKGDSSAE 360

QY 356 MSCQITATSNGEHGMNPSLOAMLMGFGDIFSMN--KAGAVMH--SGMQ-----INM- 404  
Db 361 MSVYASLFKEN-----NITGRLLLLLEEDLKDGMGIVSKGHIHFKAIEKLTHTDYINLF 415  
QY 405 --QAKONSSKTSKRGRKVMALGFSDFDLSEGDHDD-----DDGEE 446  
Db 416 HFPPLIKDSGGEPEENEKEIVNLELVFG-FHLKPGTGPQDCKMKYMEMDGD 467

RESULT 6

AAB65673

ID AAB65673 standard; Protein; 800 AA.

XX AAB65673;

DT 27-MAR-2001 (first entry)

DE Novel protein kinase, SEQ ID NO: 201.

XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;  
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;  
KW immune disorder; cardiovascular disease; neurodegenerative disease;  
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
XX inflammatory pelvic disease; multiple sclerosis; psoriasis.  
XX Homo sapiens.

XX WO200073469-A2.  
XX PN

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14842.

XX 28-MAY-1999; 99US-0136503.

XX (SUGEN-) SUGEN INC.

PI Plowman GD, Martinez R, Whyte D, Sudersanam S;

DR WPI; 2001-032161/04.

DR N-PSDB; AAF44701.

PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
PT treating immune-related diseases and disorders, cardiovascular disease,  
PT neurodegenerative diseases and/or cancers.  
XX

PS Claim 10; Fig 1; 310pp; English.

XX The present sequence is a novel protein kinase. The novel protein kinases  
CC and the nucleic acids that encode them may be used in the treatment and  
CC diagnosis of diseases associated with inappropriate kinase expression  
CC such as immune-related diseases and disorders, cardiovascular disease,  
CC neurodegenerative diseases and/or cancers. The nucleic acids and  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays. The kinase polypeptides may be used as antigens in the production  
CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
CC and kinase antagonists may also be used to down regulate kinase  
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune  
CC disorders, complications of organ transplantation, myocardial infarction,  
CC immune disorders, cardiomyopathies, strokes, renal failure,  
CC oxidative-stress related disorders, chronic inflammatory bowel disease,  
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
CC reproductive disorders.  
XX

Sequence 800 AA;

Query Match

Best Local Similarity 73.6%; Score 1748.5; DB 22; Length 800;

Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;



QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWISQDKEVAVKLLKIEKAEILSVL 60  
Db 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWISQDKEVAVKLLKIEKAEILSVL 60  
QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMHMTWATDVAKGMHY 120  
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMHMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLKICDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVIAADGVLKICDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180  
QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPRSFAELLH 240  
Db 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPRSFAELLH 240  
QY 241 QCWEADAKRPPSFQOIIISLESMSNDTSLPDKCNFHNKAEWRCEIEATLERLKKLERD 300  
Db 241 QCWEADAKRPPSFQOIIISLESMSNDTSLPDKCNFHNKAEWRCEIEATLERLKKLERD 300  
QY 301 LSPKQELKERERRLKMEQKLTQSNTPLL-LPLAARMSSESYF--ESKTEESNSAE 355  
Db 301 LSPKQELKERERRLKMEQKLTQSNTPLLPSPFEIGAWTEDDVYCWVQOLVRKGDSSAE 360  
QY 356 MSCQITATSNNGEGHGNPSTQAMLMNGFGDIFSMN--KAGAVNH--SGMQ-----INM- 404  
Db 361 MSVVASLPFKN-----NITGKRLLLEEDLDKMGIVSKGHIHFKAIEKLTHTDYINLF 415  
QY 405 ---QAKONSSKTSKRGGKVVNMLGFSDFDLSEGDDDD-----DDGEE 446  
Db 416 HFPPLIKDSGEPBENEKIVNLELVFG-FHLKPGTGPODCKWKMVMEMDGE 467

RESULT 7  
AAG75571  
ID AAG75571 standard; Protein; 349 AA.

AC AAG75571;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen protein SEQ ID NO:6335.

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma.

OS Homo sapiens.

XX WO200122920-A2.

PD 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR MPI: 2001-235357/24.

DR N-PSDB; AAH34976.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers -  
XX  
XX Claim 11; Page 7789-7790; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene

CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing  
CC inactive proteins or to supplement the patient's own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated Ps,  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
CC and AAB77789 represent sequences used in the exemplification of the  
CC present invention.

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 349 AA;

QY Query Match 59.3%; Score 1409; DB 22; Length 349;

Best Local Similarity 100.0%; Pred. No. 2.5e-119;

Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWISQDKEVAVKLLKIEKAEILSVL 60

Db 57 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWISQDKEVAVKLLKIEKAEILSVL 116

QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMHMTWATDVAKGMHY 120

Db 117 SHRNIIOFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMHMTWATDVAKGMHY 176

QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLKICDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180

Db 177 LHMEAPVKVIHRDLKSRNVIAADGVLKICDFGASRFHNHTTHMSLVGTFPMAPEVIQS 236

QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPRSFAELLH 240

Db 237 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPRSFAELLH 296

QY 241 QCWEADAKRPPSFQOIIISLESMSNDTSL 269

Db 297 QCWEADAKRPPSFQOIIISLESMSNDTSL 325

RESULT 8  
AAG03583  
ID AAG03583 standard; Protein; 141 AA.

AC AAG03583;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 7664.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST ) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX MPI: 2000-500381/45.

DR N-PSDB; AAC03589.

XX

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
PS Claim 13; SEQ ID 7664; 71bp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number  
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
CC different tissues. EST sequences usually correspond mainly to the 3'  
CC untranslated region (UTR) of the mRNA because they are often obtained  
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
CC those cases where longer cDNA sequences have been obtained, the full 5'  
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
CC ends and can therefore be used to obtain full length cDNAs and genomic  
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
CC chromosome mapping procedures. They are used to obtain upstream  
CC regulatory sequences and to design expression and secretion vectors.

Sequence 141 AA;

Query Match 31.3%; Score 744; DB 21; Length 141;  
Best Local Similarity 98.6%; Pred. No. 1.6e-59;  
Matches 139; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 102 MDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLCDFGASRFHNHT 161  
Db 1 MDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLCDFGASRFHNHT 60  
QY 162 THMSLVGTFPMPMAPEVIOQLPVSETCDTYSYGVVLMEMLTREVPFKGLEQLQVAVLVVEX 221  
Db 61 THMSLVGTFPMPMAPEVIOQLPVSETCDTYSYGVVLMEMLTREVPFKGLEQLQVAVLVVEX 120  
QY 222 NERLTISSCPRSPFAELLHQC 242  
Db 121 NERLTISSCPRSPFAELLHQC 141

RESULT 9  
AAB65552  
ID AAB65552 standard; Protein; 124 AA.  
XX  
AC AAB65552;  
XX  
DT 27-MAR-2001 (first entry)  
XX

C-terminus specific to novel human protein kinase MLK4B.

KW Human; protein kinase; antiarthritis; antisclerotic; immunosuppressive;  
KW cardiac; renal; antiinflammatory; antiasthmatic; osteopathic;  
KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;  
KW immune disorder; cardiovascular disease; neurodegenerative disease;  
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.  
XX  
OS Homo sapiens.  
XX  
PN WO200073469-A2.  
XX  
PD 07-DEC-2000.  
XX  
PF 26-MAY-2000; 2000WO-US14842.  
XX  
PR 28-MAY-1999; 99US-0136503.  
XX

PA (SUGEN-) SUGEN INC.

PI Plowman GD, Martinez R, Whyte D, Sudersanam S;  
XX  
XX WPI; 2001-032161/04.

PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and

PT treating immune-related diseases and disorders, cardiovascular disease,  
PT neurodegenerative diseases and/or cancers -  
XX  
XX  
PS Example 1; Page 116; 310pp; English.

CC The present sequence is given in a specification relating to novel  
CC protein kinases. The protein kinases and the nucleic acids that encode  
CC them may be used in the treatment and diagnosis of diseases  
CC associated with inappropriate kinase expression such as immune-related  
CC diseases and disorders, cardiovascular disease, neurodegenerative  
CC diseases and/or cancers. The nucleic acids and complementary sequences  
CC may also be used as DNA probes in diagnostic assays. The kinase  
CC polypeptides may be used as antigens in the production of antibodies of  
CC kinase expression and activity. Anti-kinase antibodies and kinase  
CC antagonists may also be used to down regulate kinase expression and  
CC activity. Diseases related to kinase expression and activity include  
CC rheumatoid arthritis, atherosclerosis, autoimmune disorders,  
CC complications of organ transplantation, myocardial infarction, immune  
CC disorders, cardiomyopathies, strokes, renal failure, oxidative-stress  
CC related disorders, chronic inflammatory bowel disease, chronic  
CC inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,  
CC psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive  
CC disorders.

Sequence 124 AA;

Query Match 26.9%; Score 638; DB 22; Length 124;  
Best Local Similarity 100.0%; Pred. No. 5.6e-50;  
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 LPLARMSEESYFESKTEESNSAEMSCQITATSNGECHGNPISLQAMLMFGDIFSNMK 391  
Db 1 LPLARMSEESYFESKTEESNSAEMSCQITATSNGECHGNPISLQAMLMFGDIFSNMK 60  
QY 392 AGAVMHSGMQINMQAKONSSKTSKRKKVNMALGFSDFDSEGGDDDDGEEBNDM 451  
Db 61 AGAVMHSGMQINMQAKONSSKTSKRKKVNMALGFSDFDSEGGDDDDGEEBNDM 120

QY 452 DNSE 455  
Db 121 DNSE 124

RESULT 10  
ABP61000  
ID ABP61000 standard; Protein; 1021 AA.  
XX  
AC ABP61000;  
XX  
DT 10-SEP-2002 (first entry)  
XX  
DE Novel human protein. SEQ ID 87.

KW Human; cytostatic; vulnerary; antiarteriosclerotic; antiparkinsonian;  
KW nootropic; neuroprotective; immunosuppressive; haemostatic;  
KW antiinflammatory; cardiac; antitumor; virucide; antithyroid;  
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;  
KW wound healing disorders; atherosclerosis; Parkinson's disease;  
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;  
KW inflammation; neoplastic disease; nervous system disorder;  
KW cardiovascular disorders; pancreatitis; respiratory disorder;  
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;  
KW developmental abnormality; gastrointestinal ulceration; neuropathy;  
KW haematological disease; metabolic disease; sperm dysfunction;  
KW thyroid disorder; hypothyroidism; brain damage; colitis;  
KW cone photo- transduction deficiency; neurological disease; stroke;  
KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;  
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;  
KW growth abnormality; precocious puberty.

Homo sapiens.

OS  
XX  
XX  
PN WO200250105-A1.

XX 27-JUN-2002.  
PD 17-DEC-2001; 2001WO-US49232.  
XX 19-DEC-2000; 2000US-256710P.  
PR 20-DEC-2000; 2000US-257048P.  
PR 09-JAN-2001; 2001US-260482P.  
PR 30-JAN-2001; 2001US-264922P.  
PR 06-FEB-2001; 2001US-266797P.  
PR 19-MAR-2001; 2001US-276988P.  
PR 04-APR-2001; 2001US-281535P.  
PR 08-MAY-2001; 2001US-289622P.  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
PA (GLAX ) GLAXO GROUP LTD.  
PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y,  
Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;  
DR WPI; 2002-508784/54.  
N-PSDB; ABQ86165.  
XX Secreted proteins and polynucleotides useful as vaccines for preventing  
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,  
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune  
PT disorder -  
XX Claim 1(a); Page 307-309; 335pp; English.  
XX The invention relates to an isolated polypeptide with signal sequences  
CC which allow it to be secreted extracellularly or membrane associated.  
CC The activity of polypeptides of the invention may be described as,  
CC cytostatic, vulnerary, antiarteriosclerotic, antiparkinsonian, nootropic,  
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,  
CC cardiant, antitumor, virucide, antithyroid, cerebroprotective, anorectic,  
CC and metabolic. Polypeptides and polynucleotides of the invention are  
CC useful in the treatment, or as a vaccine in the prevention of, cancer,  
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease  
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,  
CC inflammation, neoplastic diseases, nervous system related disorders and  
CC cardiovascular disorders, pancreatitis, respiratory disorder,  
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,  
CC developmental abnormality, gastrointestinal ulceration, neuropathy,  
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid  
CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-  
CC transduction deficiency, neurological diseases, stroke, angiogenesis,  
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,  
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,  
CC growth abnormalities, and alleviation of precocious puberty. The  
CC sequences given in records ABP60965-ABP61019 represent novel human  
CC proteins of the invention.  
XX Sequence 1021 AA;  
SQ  
Query Match 25.4%; Score 604; DB 23; Length 1021;  
Best local Similarity 34.4%; Pred. No. 1.7e-45;  
Matches 151; Conservative 74; Mismatches 142; Indels 72; Gaps 10;  
OY 6 ASFYQIKFDLQFFENC GGSFGSYRAK WISQDKEVA VVKLLK-----IEKE 53  
DB 114 SSPVHVAFERLELKE LIGAGGFGVYRATW--QGQEVAVKARQDPEDQAAAAAESVRR 171  
OY 54 AEILSVLSHRNITQFYGVILPEPNYGVTEYASLSGLYDYNRSSEMDMDHIM-TWAT 112  
DB 172 ARLFAMLRHPNITELRGVCLQOPHLCLVLEFARGALNRALARRIP---PHLVVWAV 227  
OY 113 DVAKGMHYLHMEAPYKVIHRDLKSRNVV-----AADGVLKICDFGASRFHNHTTM 164  
DB 228 QIARGMLYLHEAFVPIILHRDLKSSNILLKEIEHDDICNKTLLKTD FGLAREWHRTTKM 287  
OY 165 SLVGTPTFPWMAPEVIQSLPVSETCTDTSYGVVIMEMLTREVPFKGLEGLQVAVLVVEKNER 224

DB 288 STAGTYAMMAPEVIKSLFSKSGSDIWSYGVLLWELLTGVEVYRGIDGLAVAYGVNKL 347  
OY 225 LTIPSSCPRSFAELLHQCEADAKRPSFKOISILESMSND--TSLPDKNSFLHNKAE 282  
DB 348 LPIPTCPEPFALMKECWQODPHIRPSFALLLEQLTAIEGAVMTMPQE--SFHMODD 405  
OY 283 WRCEIATLERLKLRLDLSFKEQEL-----KERERLKMWEQKLTQ----- 325  
DB 406 WKLEIQMFDELFRKEKELRSREBELTPRALQOKSQEELLKRREQLAEREIDVLERLN 465  
OY 326 -----SNTPIILPLAARMSEESYFESKTESNSAEM---SCQIT 361  
DB 466 ILIFQLNQEKPKYKRGKFKRSRLKDKGHRISLPSDFQHKITVOASPNLDKRRSLNS 525  
OY 362 ATSNGEHGMNPSLOAMTL 380  
DB 526 SSSPPSSPTMPRLRAIQL 544

RESULT 11  
AAB85513  
ID AAB85513 standard; protein; 719 AA.  
XX AAB85513;  
AC AAB85513;  
XX 25-SEP-2001 (first entry)  
DT Human protein kinase SGK067.  
DE  
XX Protein kinase; enzyme; cytostatic; nootropic; neuroprotective; human;  
KW antiparkinsonian; virucide; antibacterial; antifungal; antimigraine;  
KW analgesic; hypotensive; hypertensive; immunosuppressive; antiallergic;  
KW antipsoriatic; antirheumatic; antiarthritic; ophthalmological; anorectic;  
KW osteopathic; thrombolytic; antiarteriosclerotic; antiasthmatic;  
KW vasotropic; antidiabetic; gene therapy.  
OS Homo sapiens.  
XX WO200155356-A2.  
PN 02-AUG-2001.  
PD 25-JAN-2001; 2001WO-US02337.  
PF 25-JAN-2000; 2000US-0178078.  
XX 31-JAN-2000; 2000US-0179364.  
PR 17-FEB-2000; 2000US-0183173.  
PR 17-MAR-2000; 2000US-0190162.  
PR 29-MAR-2000; 2000US-0193404.  
PR 13-NOV-2000; 2000US-0247013.  
XX (SUGE-) SUGEN INC.  
PA Plowman G, Whyte D, Manning G, Sudarsanam S, Martinez R;  
PI WPI; 2001-476202/51.  
DR N-PSDB; AAH46913.  
XX kinase polypeptides useful for treating cancers, Alzheimer's disease,  
PT viral infections, diabetes, obesity, organ transplant rejection and  
PT rheumatoid arthritis -  
PT Claim 7, Page 217; 218pp; English.  
XX The invention provides human protein kinases and protein kinase-like  
CC enzymes and polynucleotides encoding the polypeptides. The kinase  
CC polypeptides and their modulators are useful for treating a disease or  
CC disorder such as cancer, immune-related diseases, cardiovascular disease,  
CC brain or neuronal-associated disease and metabolic disorders, including  
CC cancers of tissues, cancers of hematopoietic origin, diseases of the  
CC central nervous system, diseases of the peripheral nervous system,  
CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic

CC lateral sclerosis, viral infections, infections caused by prions,  
CC bacteria and fungi, ocular diseases, migraines, pain, sexual dysfunction,  
CC mood disorders, attention disorders, cognition disorders, hypotension,  
CC hypertension, psychotic disorders, neurological disorders, dyskinesias,  
CC metabolic disorders, and organ transplant rejection. They are also useful  
CC for treating rhinitis, autoimmunity, atherosclerosis, psoriasis,  
CC osteoarthritis, asthma, chronic inflammatory pelvic disease, chronic  
CC inflammatory bowel disease, rheumatoid arthritis, metabolic disorders  
CC such as diabetes, obesity, cardiovascular diseases such as reperfusion  
CC injury, coronary thrombosis, clotting disorders and atherosclerosis,  
CC ocular diseases such as glaucoma, retinopathy and macular degeneration,  
CC psychiatric and neurological disorders such as anxiety, schizophrenia,  
CC dementia, manic depression, etc. The polynucleotides are useful in gene  
CC therapy techniques to treat the above mentioned disorders. Sequences  
CC AAB85491-85522 represent the human protein kinases of the invention.

XX Sequence 719 AA;

SO Query Match 25.2%; Score 599.5; DB 22; Length 719;  
Best Local Similarity 33.3%; Pred. No. 2.6e-45;  
Matches 150; Conservative 75; Mismatches 146; Indels 79; Gaps 9;

QY 6 ASFVQIKFDLQFFENCSSGSGSVYRAKWSQDEKAVAKLLK-----IEKE 53  
DB 114 SSPVHVAFERLEKELIGAGGFGQVYRATW--QGQEVAVKARQDPEDAAAAAESVRR 171  
QY 54 AEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSLYDI-----NSNRSEE 101  
DB 172 ARLFAMLRHPNIIELRGVCLQOPHLCLVLEFARGALNRALAAANAPDPRAPRRARR 231  
QY 102 MDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVI-----AADGLKICDFG 153  
DB 232 IPRHVLVWVAQIARGMLYLHBEAFVPIHRDLKSSNILLKEIEHDDICNKTLCITDFG 291  
QY 154 ASRFHNHTTHMSLVGTFPWWAPREVIOSLPVSETCDTYSYGVVLMWMLTREVPFKGLEGLQ 213  
DB 292 LAREWHRTTKMSTAGTYAMAPREVIKSSLSFGSGSDIWSYGVLLWELLTGEVPRGIDGLA 351  
QY 214 VAWLVVEKNERLTIPSSCPSPFAELLHQCEADAKKRPSEKQIISLESMSND--TSLPD 271  
DB 352 VAYGVAVNKLTLPSTCPPEFAKLMKECWQODPHIRPSFALILEQLTAIEGAVWTEMPQ 411  
QY 272 KNSFLHNKAEMRCIEATLERLKLRLDSFKEQEL-----KERERLKMWEQKLT 324  
DB 412 E--SFHSMQDDWKLEIQOMFDELRTKEKELRSREELTRAALQOKSQEELLKRREQQLAE 469  
QY 325 Q-----SNTPLLLPLAARMESESYFESKTEESNSA 354  
DB 470 REIDVLERELNILIFQLNQEKPKVKRKGKFKRSRLKLDGHRISLPSDFQHKITVQASP 529  
QY 355 EM---SCQITATSNGEHGMNPSLOAMML 380  
DB 530 NLDKRRSLNSSSSSPSSPTMPRLRAIQL 559

RESULT 12  
AAB80923  
ID AAB80923 standard; Protein; 1036 AA.

XX AAB80923;  
DT 08-OCT-2002 (first entry)  
XX Novel human protein (NHP) kinase.  
XX Novel human protein, NHP; kinase; human; enzyme.  
OS Homo sapiens.

XX Key Location/Qualifiers  
FH Misc-difference 925 /note= "encoded by WGT"  
FT  
XX

PN WO200255685-A2.  
XX  
XX 18-JUL-2002.  
PD  
XX 10-DEC-2001; 2001WO-US47606.  
PF  
XX 11-DEC-2000; 2000US-254744P.  
PR  
XX (LEXI-) LEXICON GENETICS INC.  
PA  
XX Hu Y, Kieke JA, Donoho G;  
PI  
XX WPI; 2002-566739/60.  
DR  
XX N-PSDB; ABN86357, ABN86358.

PT Novel human kinase polynucleotide encoding a protein that shares  
PT structural similarity with animal kinases for therapeutic, diagnostic  
PT and pharmacogenomic applications.  
PS Claim 1, Page 37-39; 41pp; English.

CC The invention relates to a novel human protein (NHP), kinase that shares  
CC structural similarity with animal kinases. The kinase polynucleotides are  
CC useful in therapeutic, diagnostic and pharmacogenomic applications and  
CC for identifying compounds that modulate, i.e. act as agonists or  
CC antagonists of the gene expression or gene product activity. The present  
CC sequence represents the NHP kinase.

XX Sequence 1036 AA;

SO Query Match 25.2%; Score 599.5; DB 23; Length 1036;  
Best Local Similarity 33.3%; Pred. No. 4.4e-45;  
Matches 150; Conservative 75; Mismatches 146; Indels 79; Gaps 9;

QY 6 ASFVQIKFDLQFFENCSSGSGSVYRAKWSQDEKAVAKLLK-----IEKE 53  
DB 114 SSPVHVAFERLEKELIGAGGFGQVYRATW--QGQEVAVKARQDPEDAAAAAESVRR 171  
QY 54 AEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSLYDI-----NSNRSEE 101  
DB 172 ARLFAMLRHPNIIELRGVCLQOPHLCLVLEFARGALNRALAAANAPDPRAPRRARR 231  
QY 102 MDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVI-----AADGLKICDFG 153  
DB 232 IPRHVLVWVAQIARGMLYLHBEAFVPIHRDLKSSNILLKEIEHDDICNKTLCITDFG 291  
QY 272 KNSFLHNKAEMRCIEATLERLKLRLDSFKEQEL-----KERERLKMWEQKLT 324  
DB 412 E--SFHSMQDDWKLEIQOMFDELRTKEKELRSREELTRAALQOKSQEELLKRREQQLAE 469  
QY 325 Q-----SNTPLLLPLAARMESESYFESKTEESNSA 354  
DB 470 REIDVLERELNILIFQLNQEKPKVKRKGKFKRSRLKLDGHRISLPSDFQHKITVQASP 529  
QY 355 EM---SCQITATSNGEHGMNPSLOAMML 380  
DB 530 NLDKRRSLNSSSSSPSSPTMPRLRAIQL 559

RESULT 13  
AAE21717  
ID AAE21717 standard; Protein; 1097 AA.

XX AAE21717;  
AC  
XX



| DT | 16-JUL-2002   | (first entry)                             |
|----|---|---|
| XX | Human PKIN-12 protein.  |   |
| XX | Human; kinase; enzyme; PKIN-12 protein; immune system disorder; anaemia; acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease; asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease; AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder; Down's syndrome; gene therapy; protein therapy; cytostatic. |   |
| OS | Homo sapiens.   |   |
| XX |   |   |
| XX |   |   |
| FH | Key   | Location/Qualifiers                       |
| FT | Peptide   | 1..17                                     |
| FT |   | /label= Signal_peptide                    |
| FT | Protein   | 18..1097                                  |
| FT |   | /note= "Mature human PKIN-12 protein"     |
| FT |   | 55..114                                   |
| FT | Domain  | /note= "SH3 domain"                       |
|    |   | 144..403                                  |
|    | Domain  | /note= "Eukaryotic protein kinase domain" |
| FT |   | 146..396                                  |
| FT | Domain  | /note= "Protein kinase domain"            |
| FT |   | 163..396                                  |
| FT | Domain  | /note= "Protein kinase domain"            |
| FT |   | 220..233                                  |
| FT | Domain  | /note= "Tyrosine kinase catalytic domain" |
| FT |   | 258..276                                  |
| FT | Domain  | /note= "Tyrosine kinase catalytic domain" |
| FT |   | 311..321                                  |
| FT | Domain  | /note= "Tyrosine kinase catalytic domain" |
| FT |   | 330..352                                  |
| FT | Domain  | /note= "Tyrosine kinase catalytic domain" |
| FT |   | 374..396                                  |
| FT | Domain  | /note= "Tyrosine kinase catalytic domain" |
| FT |   | 438..749                                  |
| FT | Domain  | /note= "Leucine zipper domain"            |
| FT |   | 869..893                                  |
| FT | Domain  | /note= "Leucine zipper domain"            |
| XX |   |   |
| XX | WO200218557-A2.   |   |
| PN |   |   |
| PD | 07-MAR-2002.  |   |
| XX |   |   |
| XX |   |   |
| PF | 31-AUG-2001; 2001WO-US27219.  |   |
| XX |   |   |
|    | 31-AUG-2000; 2000US-229873P.  |   |
|    | 08-SEP-2000; 2000US-231357P.  |   |
|    | 14-SEP-2000; 2000US-232654P.  |   |
|    | 22-SEP-2000; 2000US-234902P.  |   |
| PR | 29-SEP-2000; 2000US-236499P.  |   |
| PR | 06-OCT-2000; 2000US-238389P.  |   |
| PR | 13-OCT-2000; 2000US-240542P.  |   |
| XX |   |   |
| PA | (INCY-) INCYTE GENOMICS INC.  |   |
| XX |   |   |
| PI | Bandman O, Nguyen DB, Walia NK, Hafalia AJA, Yao MG, Gandhi AR;   |   |
| PI | Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;   |   |
| PI | Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YT;   |   |
| PI | Azimzai Y, Burrill JD, Marcus GA, Zingler KA, Lu DAM, Lal PG;   |   |
| PI | Ramkumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;   |   |
| XX | Burford N;  |   |
| DR | WPI; 2002-329769/36.  |   |
| DR | N-PSDB; AAD34309.   |   |
| XX |   |   |
| XX | New human kinases, useful for diagnosing, treating or preventing immune   |   |
| PT | system disorders (e.g. Crohn's disease), neurological disorders (e.g.   |   |
| PT | epilepsy), or cell proliferative disorders (e.g. cancers such as  |   |
| PT | leukemia or lymphoma)   |   |
| XX |   |   |
| PS | Claim 67; Page 171-173; 218pp; English.   |   |

[illegible]

FT Region 154..207  
/note= "Receptor tyrosine kinase"  
FT Region 181..228  
/note= "Receptor tyrosine kinase"  
FT Region 210..223  
/note= "Tyrosine kinase catalytic site"  
FT Region 232..254  
/note= "Receptor tyrosine kinase"  
FT Region 248..266  
/note= "Tyrosine kinase catalytic site"  
FT Region 290..337  
/note= "Receptor tyrosine kinase"  
FT Region 291..340  
/note= "Receptor tyrosine kinase"  
FT Region 298..330  
/note= "Receptor tyrosine kinase"  
FT Region 301..311  
/note= "Tyrosine kinase catalytic site"  
FT Region 320..342  
/note= "Tyrosine kinase catalytic site"  
FT Region 337..389  
/note= "Receptor tyrosine kinase"  
FT Region 345..389  
/note= "Receptor tyrosine kinase"  
FT Region 356..404  
/note= "Receptor tyrosine kinase"  
FT Region 364..386  
/note= "Tyrosine kinase catalytic site"  
PN WO200181555-A2.  
PD 01-NOV-2001.  
XX 20-APR-2001; 2001WO-US12992.  
PF 20-APR-2000; 2000US-199021P.  
PR 28-APR-2000; 2000US-200226P.  
PR 05-MAY-2000; 2000US-202339P.  
PR 11-MAY-2000; 2000US-203505P.  
PR 18-MAY-2000; 2000US-205564P.  
PR 26-MAY-2000; 2000US-207739P.  
PR 01-JUN-2000; 2000US-208795P.  
PA (INCY-) INCYTE GENOMICS INC.  
PI Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;  
PI Bandman O, Lu DAM, Lal P, Burford N, Khan PA, Walla NK, Yao MG;  
PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;  
PI Walsh RT, Ramkumar J, Borowsky ML, Au-young J, Hillman JL;  
PI Gururajan R;  
DR MPI; 2001-611740/70.  
DR N-PSDB; AAD18824.  
XX Human kinases and nucleic acids, useful for preventing diagnosing and  
XX treating cancers, inflammation and immune disorders -  
PS Claim 1; Page 134-136; 166pp; English.  
XX The present invention relates to human kinases (PKIN) and the nucleic  
XX acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is  
XX used in the prevention, diagnosis and treatment of diseases cancers,  
XX adenocarcinoma, leukemia, sarcoma, immune disorder, Addison's disease,  
XX acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,  
XX gout, microbial infections, cardiovascular disease and/or inflammation,  
XX myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial  
XX infarction, catarract, growth and development disorder, seizure disorder,  
XX pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage  
XX disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.  
XX PKIN may be used to treat disorders associated with decreased PKIN  
XX expression by rectifying mutations or deletions in a patient's genome  
XX that affect the activity of PKIN by expressing inactive proteins or to

CC supplement the patients own production of PKIN. PKIN nucleic acids may be  
CC used to produce the PKIN polypeptide, by inserting the nucleic acids into  
CC a host cell and culturing the cell to express the protein. PKIN nucleic  
CC acid and its complementary sequences may also be used as DNA probes in  
CC diagnostic assays to detect and quantitate the presence of similar  
CC nucleic acid sequences in samples and therefore which patients may be  
CC in need of restorative therapy. The present sequence is human PKIN-9  
CC protein.  
XX SQ Sequence 1046 AA;  
Query Match 24.8%; Score 589; DB 22; Length 1046;  
Best Local Similarity 39.1%; Pred. No. 4.1e-44;  
Matches 134; Conservative 59; Mismatches 118; Indels 32; Gaps 5;  
QY 10 QIKFDDLOFFENCGGSGSVYRAKWSQDEVAVK-----KLLKIEKEAETL 57  
DB 128 EIDFAELLTBEIIGIGFGKVYRAFWIGD--EVAVKARHPDDEDISQTIENVROEAKLF 185  
QY 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEMDMDHMTWATDVAKG 117  
DB 186 AMLKHPNIIALRGVCLKEPNLCLVMEFARGPPLNRVLSGKR--IPDILVNMAVQIARG 242  
QY 118 MEYLMAPVVKVTHDLKSRNVI-----AADGLKICDFGASRFHNHTHMSLVGT 169  
DB 243 MNVLLDEAIVPIHRDLKSSNIILOXYENGDSLKNILKITDFGLAREWHRTTKMSAAGT 302  
QY 170 FPMWAPVIOQLPVSETCDTYSYGVVLEWMLTREVPFKEGLQVAVLVVEKNERLTIPS 229  
DB 303 YAMWAPVIRASMSFSGSDWSYGVLLWELLTGEVPPFGIDGLAVAYGVAANKLALPIPS 362  
QY 230 SCPRSFALLHOCWEADAKKRPSPFKQIISLESMSNDTSLPDKNSFLHNKAENRCEIEA 289  
DB 363 TCPEPFAKLMEDCWNPDPHSRPSFTNILDQLTTIEBSGFPEMPKDSFHLQDNWKEIOE 422  
QY 290 TLERLKKLRDLSFKEQEL-----KERERLKMWEQKLTQ 325  
DB 423 MFDOLRAKEKELRTWEEELTPRALQCKNQEELLRRRQELAE 465  
RESULT 15  
AAE22763  
ID AAE22763 standard; Protein; 847 AA.  
XX AAE22763;  
AC AAE22763;  
XX 09-AUG-2002 (first entry)  
DT Human mitogen activated protein kinase, MAP3K11.  
XX Human mitogen activated protein kinase, MAP3K11.  
DE Human; cytosolic; antisense gene therapy; screening; protein kinase;  
KW cancer; liver; colon; tumour; inflammation; arthritic synovium; MAP3K11;  
KW enzyme; mitogen activated protein kinase.  
XX Homo sapiens.  
OS WO200224947-A2.  
PN 28-MAR-2002.  
PD 20-SEP-2001; 2001WO-IB02237.  
PF 20-SEP-2000; 2000US-233999P.  
PR 02-OCT-2000; 2000US-237419P.  
PR 02-OCT-2000; 2000US-237423P.  
PR 04-OCT-2000; 2000US-238558P.  
PR 10-MAY-2001; 2001US-290555P.  
XX (KINE-) KINETEK PHARM INC.  
PA (UYBR-) UNIV BRITISH COLUMBIA.  
XX Yoganathan T, Delaney AD;

DR WPI; 2002-394145/42.  
DR N-PSDB; AAD36139.

DR N-PSDB; AAD36139.

XX Diagnosing cancer, comprises determining the upregulation of expression  
PT of a nucleic acid sequence encoding a protein kinase or upregulation of  
PT expression of the protein kinase, in the cancer -  
PT

xx Claim 1; Page 60-62; 87pp; English.  
ps

XX The invention relates to a method for screening biologically active agent  
CC that modulates cancer associated protein kinase function. The invention  
CC also relates to a method for diagnosing cancer comprising determining the  
CC upregulation of expression of a nucleic acid sequence encoding a protein  
CC kinase. The method is useful for diagnosing cancer. A protein kinase is  
CC useful for screening biological agents that modulate cancer associated  
CC protein kinase function. Downregulating the activity of protein kinase is  
CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon  
CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy  
CC derived tumours and inflammatory samples such as arthritic synovium, for  
CC amplified DNA in the cell or increased expression of corresponding mRNA  
CC or protein and is also useful to detect differences in expression levels  
CC such as molecular weight, amino acid and nucleotide sequences between the  
CC two cells. The present sequence is human mitogen activated protein  
CC kinase, MAP3K11.

22

20

XX

| SQ | Sequence | 847 | AA: |
|----|----------|-----|-----|
|----|----------|-----|-----|

|                       |                  |                    |            |             |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match           | 24.5%;           | Score 581.5;       | DB 23;     | Length 847; |
| Best Local Similarity | 37.4%;           | Pred. No. 1.4e-43; |            |             |
| Matches 132;          | Conservative 60; | Mismatches 122;    | Indels 39; | Gaps 7;     |

13 FDDLOFFENCGGSGFSGVYRAKWISQ-----DKEVAVKLLIKIEKEAELSTLS 61

Db 114 FQELPLEEVIGIGFGKVTGRSGNELVAVKARQDPDEDISV-TAESVRQEARLLFAMLA 1/2

62 H R N I Q F Y G V I L E P P N Y G I V T E A S L G S L Y D I N S N R S E E M D M H I M - T W A T D V A K G M Y 120

Db 173 HPNITALKAVCTLEEPNCLVMEYAGCDLSRALAGRVP---PHYLVNWAQIARGMHY 228

121 LHMEAPVKYIHRDLKSRNV-----IAADGV-----LKICDFGASRFHHHTTMSLVGTFPW 177

Db 229 LHCEALVPYIHRDKSNNILLLQIESDDMEHKLKITDFGLAREWHTTQMSAAGTYAW 28

QY 173 MAPEVIOQLPYSETCDTYSYGVLMWMLTREYPFKLEGLQVAMLVVEKNERLITIPSSCP 23

```
Db      289 MAPEVIKASTFSGSDVMSFGVLLMELLTGEVPYRGIDCLAVAYGVAUNKLTPIPSTCP 34
```

233 RSFAELHQWEADAKRPSFKQISILESMSNDTSLPDKNSFLHNKAEMCEIATLE 29

349 EPFAQLMADCWAQDPHRRPDDFASILQQLALEAQLREMPRDSFHSMQEGWKREIQGLFD 40

QY 293 RLKLELDSEKOE-----LKERERRLKMEQKXTEQSNTPLL 331

Db 409 ELRAKEKELLISREELTRAPAEORSQAEQLRREHLAQWELVFERELTLLL 461

Search completed: May 1, 2003, 20:35:33  
Job time : 92.5732 secs

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:38:08 ; Search time 29.1911 Seconds  
(without alignments)  
1344.947 Million cell updates/sec

Title: US-09-757-982-5  
Perfect score: 2375  
Sequence: 1 MSSLGASFVQIKFDDLOFFE.....GDDDDDDGEEEDNDMDNSE 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues  
Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

|     |  |
|-----|--|
| 1:  | /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*  |
| 2:  | /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*   |
| 3:  | /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*  |
| 4:  | /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*  |
| 5:  | /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*  |
| 6:  | /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*  |
| 7:  | /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:* |
| 8:  | /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*  |
| 9:  | /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*  |
| 10: | /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*  |
| 11: | /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*  |
| 12: | /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*  |
| 13: | /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*  |
| 14: | /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*  |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID                    | Description        |
|------------|-------|-------------|-----------|-----------------------|--------------------|
| 1          | 2375  | 100.0       | 455       | 10 US-09-757-982-5    | Sequence 5, Appli  |
| 2          | 599.5 | 25.2        | 1036      | 12 US-10-014-882-2    | Sequence 2, Appli  |
| 3          | 593   | 25.0        | 394       | 10 US-09-862-027-19   | Sequence 19, Appli |
| 4          | 581.5 | 24.5        | 847       | 9 US-10-143-133-2     | Sequence 2, Appli  |
| 5          | 568.5 | 23.9        | 966       | 10 US-09-771-161A-197 | Sequence 197, App  |
| 6          | 545.5 | 23.0        | 328       | 10 US-09-862-027-18   | Sequence 18, Appli |
| 7          | 504.5 | 21.2        | 850       | 10 US-09-904-389-2    | Sequence 2, Appli  |
| 8          | 487   | 20.5        | 263       | 10 US-09-840-704-5    | Sequence 5, Appli  |
| 9          | 481   | 20.3        | 579       | 9 US-10-158-895-4     | Sequence 4, Appli  |
| 10         | 481   | 20.3        | 590       | 9 US-10-158-895-15    | Sequence 15, Appli |
| 11         | 424   | 17.9        | 92        | 9 US-09-764-868-799   | Sequence 799, App  |
| 12         | 392.5 | 16.5        | 835       | 10 US-09-947-199-8    | Sequence 8, Appli  |
| 13         | 385   | 16.2        | 1036      | 10 US-09-771-161A-255 | Sequence 255, App  |
| 14         | 385   | 16.2        | 1036      | 10 US-09-771-161A-256 | Sequence 256, App  |
| 15         | 382.5 | 16.1        | 251       | 8 US-08-987-689A-32   | Sequence 32, Appli |
| 16         | 379.5 | 16.0        | 505       | 9 US-09-977-260-6     | Sequence 6, Appli  |
| 17         | 379.5 | 16.0        | 505       | 9 US-09-977-261-6     | Sequence 6, Appli  |
| 18         | 379.5 | 16.0        | 505       | 10 US-09-977-269-6    | Sequence 6, Appli  |
| 19         | 379.5 | 16.0        | 505       | 10 US-09-982-610-20   | Sequence 20, Appli |

|    |       |      |      |                       |                    |
|----|-------|------|------|-----------------------|--------------------|
| 20 | 377   | 15.9 | 835  | 10 US-09-947-199-2    | Sequence 2, Appli  |
| 21 | 376.5 | 15.9 | 537  | 9 US-09-977-260-11    | Sequence 11, Appli |
| 22 | 376.5 | 15.9 | 537  | 9 US-09-977-261-11    | Sequence 11, Appli |
| 23 | 376.5 | 15.9 | 537  | 10 US-09-977-269-11   | Sequence 212, App  |
| 24 | 374.5 | 15.8 | 537  | 10 US-09-771-161A-212 | Sequence 213, App  |
| 25 | 374.5 | 15.8 | 537  | 10 US-09-771-161A-213 | Sequence 12, Appli |
| 26 | 365.5 | 15.4 | 536  | 9 US-09-977-260-12    | Sequence 12, Appli |
| 27 | 365.5 | 15.4 | 536  | 9 US-09-977-261-12    | Sequence 12, Appli |
| 28 | 365.5 | 15.4 | 536  | 10 US-09-977-269-12   | Sequence 3, Appli  |
| 29 | 358.5 | 15.1 | 675  | 9 US-10-186-399-3     | Sequence 4, Appli  |
| 30 | 358.5 | 15.1 | 675  | 9 US-09-977-260-4     | Sequence 4, Appli  |
| 31 | 358.5 | 15.1 | 675  | 9 US-09-977-261-4     | Sequence 4, Appli  |
| 32 | 358.5 | 15.1 | 675  | 10 US-09-977-269-4    | Sequence 714, App  |
| 33 | 358   | 15.1 | 764  | 10 US-09-925-302-714  | Sequence 29, Appli |
| 34 | 357.5 | 15.1 | 425  | 10 US-09-828-313-29   | Sequence 7, Appli  |
| 35 | 355   | 14.9 | 310  | 9 US-09-939-833-7     | Sequence 7, Appli  |
| 36 | 355   | 14.9 | 310  | 10 US-09-939-754-7    | Sequence 7, Appli  |
| 37 | 355   | 14.9 | 310  | 10 US-09-939-832-7    | Sequence 2, Appli  |
| 38 | 355   | 14.9 | 822  | 9 US-09-757-415A-2    | Sequence 13, Appli |
| 39 | 349.5 | 14.7 | 536  | 9 US-09-977-260-13    | Sequence 10, Appli |
| 40 | 349.5 | 14.7 | 536  | 9 US-09-929-266-10    | Sequence 13, Appli |
| 41 | 349.5 | 14.7 | 536  | 9 US-09-977-261-13    | Sequence 13, Appli |
| 42 | 349.5 | 14.7 | 536  | 10 US-09-977-269-13   | Sequence 2, Appli  |
| 43 | 349.5 | 14.7 | 1308 | 10 US-09-940-101-2    | Sequence 7, Appli  |
| 44 | 347.5 | 14.6 | 450  | 9 US-09-977-260-7     | Sequence 7, Appli  |
| 45 | 347.5 | 14.6 | 450  | 9 US-09-977-261-7     | Sequence 7, Appli  |

ALIGNMENTS

RESULT 1  
US-09-757-982-5  
Sequence 5, Application US/09757982  
Patent No. US20020094559A1  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OR INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: MNI-050  
CURRENT APPLICATION NUMBER: US/09/757,982  
CURRENT FILING DATE: 2001-01-10  
PRIOR APPLICATION NUMBER: 09/163,115  
PRIOR FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-757-982-5

| Query Match           | Score   | DB                     | Length               |
|-----------------------|---------|------------------------|----------------------|
| Best Local Similarity | 100.0%; | 10;                    | 455;                 |
| Matches               | 455;    | Conservative 0;        | Mismatches 0;        |
| Indels                | 0;      | Gaps                   | 0;                   |
| QY                    | 1       | MSSLGASFVQIKFDDLOFFENC | GGGFGSVYRAKMSQDKVA   |
| Db                    | 1       | MSSLGASFVQIKFDDLOFFENC | GGGFGSVYRAKMSQDKVA   |
| QY                    | 61      | SHRNIIQFYGVILEPNIYIVTE | YASLGLDYIINSNRSEEMDM |
| Db                    | 61      | SHRNIIQFYGVILEPNIYIVTE | YASLGLDYIINSNRSEEMDM |
| QY                    | 121     | LHMEAPVKVIHRDLKSRNVIA  | DGLKICDFGSRFHNHTTMS  |
| Db                    | 121     | LHMEAPVKVIHRDLKSRNVIA  | DGLKICDFGSRFHNHTTMS  |
| QY                    | 181     | LPVSETCDTYSYGVVLMEMLT  | REVPFKGLGLQYAWLVVEKN |
| Db                    | 181     | LPVSETCDTYSYGVVLMEMLT  | REVPFKGLGLQYAWLVVEKN |
| QY                    | 241     | QCWEADAKKRPSFKQIISLES  | MSNDTSLPDKCNSFLHNKA  |
| Db                    | 241     | QCWEADAKKRPSFKQIISLES  | MSNDTSLPDKCNSFLHNKA  |



Db 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIATLERLKKLERD 300  
QY 301 LSFKEQELKERERRLKQWEQKLTQSNTPLLPLAARMSESYFESKTEESNSAEMSCQI 360  
Db 301 LSFKEQELKERERRLKQWEQKLTQSNTPLLPLAARMSESYFESKTEESNSAEMSCQI 360  
QY 361 TATSNGEHGMNPSLOAMLMNGFDIFSMNKAGAVMHSQMOINMQAKONSSTTSKRGK 420  
Db 361 TATSNGEHGMNPSLOAMLMNGFDIFSMNKAGAVMHSQMOINMQAKONSSTTSKRGK 420  
QY 421 KVMNALGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455  
Db 421 KVMNALGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 2

US-10-014-882-2  
; Sequence 2, Application US/10014882  
; Patent No. US20020107384A1  
GENERAL INFORMATION:

APPLICANT: Hu, Yi  
APPLICANT: Kieke, James  
APPLICANT: Donoho, Gregory  
; TITLE OF INVENTION: No. US20020107384A1e1 Human kinase and Polynucleotides Encoding  
; FILE REFERENCE: LEX-0279-USA  
; CURRENT APPLICATION NUMBER: US/10/014,882  
; PRIOR FILING DATE: 2001-12-11  
; PRIOR APPLICATION NUMBER: US 60/254,744  
; PRIOR FILING DATE: 2000-12-11  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1036  
; TYPE: PRT  
; ORGANISM: homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(1036)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-014-882-2

Query Match 25.2%; Score 599.5; DB 12; Length 1036;  
Best Local Similarity 33.3%; Pred. No. 2.6e-36;  
Matches 150; Conservative 75; Mismatches 146; Indels 79; Gaps 9;

QY 6 ASFVQIKFDDLOFFENCGGSGSVYRAKWISQDKEVAVKLLK-----IERE 53  
Db 114 SSPVHVAFERLELKEKELIGAGFGQYRATW--QGEVAAYKARQDPEQDAAAAESEVARE 171  
54 AELISVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYI-----NSNRSEE 101  
Db 172 ARLFAMLRHPNIIELRGVCLQPHILCLVLEFARGALNRALAAANAADPRAPGRARR 231  
QY 102 MDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVY-----AADGVLCIDPF 153  
Db 232 IPRHVLVNWAVQIARGMLYLHEEAFVPIHRDLKSSNILLLEKIEHDDICNKTITDFG 291  
QY 154 ASRFHNHTTHMSLVGTFPMWAPVYIQLPVSETCDTYSYGVVLMEMLTREVFPKLEGLO 213  
Db 292 LAREWHRTTKMSTAGTYAMMAPEVIKSSLSFGSDIWSYGVLLWELLTGEVPRGIDGLA 351  
QY 214 VAMLVVEKNERLTIPSSCPRSFAELLHQWEADAKRPSFKQIISILESMSND--TSLPD 271  
Db 352 VAYGVAVNKLTLPIPTCPEPAKMKCEWCQODPHIRPSFALLLEQLTAIEGAVMTEMQ 411  
QY 272 KCSNGLHNKAEMRCEIATLERLKKLERDLSFKEQEL-----KERERLQWEQKLT 324  
Db 412 E--SFHSMQDDWKLEIQMFDELTKELRSREELTRALQCKQOEELLKRREGQALAE 469  
QY 325 Q-----SNTPLLLPLAARMSESESYFESKTEESNSA 354  
Db 470 REIDVLERELNILLIFOLNQEKPVKYKRGKFKRSRLKLDGHRISLPSDFQHKITVQASP 529

QY 355 EM-----SCQITATSNGEHGMNPSLOAMML 380  
Db 530 NIDKRRSLNSSSSSPSSPTMPRLRAIQL 559

RESULT 3

US-09-862-027-19  
; Sequence 19, Application US/09862027  
; Patent No. US20020142428A1  
GENERAL INFORMATION:

APPLICANT: Hodge, Martin R.  
; TITLE OF INVENTION: No. US20020142428A1e1 kinases and Uses Thereof  
; FILE REFERENCE: 35800/234862  
; CURRENT APPLICATION NUMBER: US/09/862,027  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: US 09/345,473  
; PRIOR FILING DATE: 1999-06-30  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-862-027-19

Query Match 25.0%; Score 593; DB 10; Length 394;  
Best Local Similarity 39.3%; Pred. No. 2.4e-36;  
Matches 133; Conservative 59; Mismatches 114; Indels 32; Gaps 5;

QY 15 DLQFFENCGGSGSVYRAKWISQDKEVAVK-----KLKIEKAEILSVLSH 62  
Db 2 ELTLEEIIGIGFGKYRAFWIGD--EVAVKARHDPDEDISQTIENVRQEAFLFAMLK 59  
QY 63 NIIQFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMDHIMTWATDVAKGMHYLH 122  
Db 60 PNITARGVCLKEPNCLVMEFARGPLNRVLSGR--IPPDILVNWAVQIARGMNYLH 116  
QY 123 MEAPVKVIHRDLKSRNVY-----AADGVLCIDFGASRFHNHTTHMSLVGTFPMA 174  
Db 117 DEAIVPIIHRDLKSSNILLIQKVENGDLSNKLKLTDFGLAREWHRTTKMSAGTYAMMA 176  
QY 175 PEVIQSLPVSETCDTYSYGVVLMEMLTREVFPKLEGLOVAVMLVVEKNERLTIPSSCPRS 234  
Db 177 PEVIRASMFSGSDVWSYGVLLWELLTGEVPRGIDGLRAYGVAAMKALPIPTCPEP 236  
QY 235 FAELLHQWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIATLERL 294  
Db 237 FAKLMEDCWNPDPHSRPSFTNILDQLTTIEESGFEMPCKDSFHLQDNWKHEIQEMFDQL 296  
QY 295 KLERDLSFKEQEL-----KERERLQWEQKLTQ 325  
Db 297 RAKEKELRTWEEELTRALQCKQOEELLRRREQELAER 334

RESULT 4

US-10-143-133-2  
; Sequence 2, Application US/10143133  
; Publication No. US20020197658A1  
GENERAL INFORMATION:

APPLICANT: Delaney, Allen  
; TITLE OF INVENTION: Cancer Associated Protein Kinase and Its Use  
; FILE REFERENCE: KINE-023  
; CURRENT APPLICATION NUMBER: US/10/143,133  
; PRIOR FILING DATE: 2002-05-09  
; PRIOR APPLICATION NUMBER: 60/290,555  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 847  
; TYPE: PRT  
; ORGANISM: Homo sapien



```

; CURRENT APPLICATION NUMBER: US/09/904,389
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/218,307
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Cucumis melo
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (154) ..(154)
; OTHER INFORMATION: Xaa = Any Amino Acid
;
US-09-904-389-2

```

|                       |                  |                    |            |             |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match           | 21.2%;           | Score 504.5;       | DB 10;     | Length 850; |
| Best Local Similarity | 40.9%;           | Pred. No. 2.3e-29; |            |             |
| Matches 112;          | Conservative 41; | Mismatches 106;    | Indels 15; | Gaps 4;     |

|    |   |     |
|----|---|-----|
| 3  | SLGASFVQIKEDDLQFFENCGGSPFSVYRAKMWISQDKAEVAVKKLKE-----K              | 52  |
| Db | 564 SLGLEDLVTPWMDLDRBKIGAGSFGTVYRGEMHGSD--VAVKILTEQDFHPERVNEFLR     | 621 |
| QY | 53 EAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLDYDINSNRSEEMDMHMTWAT       | 112 |
| Db | 622 EVAIMKSLRHNPVLFEMGAVTKPPLNSTLYTEYLSRGLYRLHKSQVKDIDETRINMAF      | 681 |
| QY | 113 DVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLKICDFGASRPHNHT--THMSLVGTF    | 170 |
| Db | 682 DVAKGMNYLHRDP-PIVHRDLKSPNLVYDKKYTVKVCDFGLSRUKARTFLSSKSAAGTP     | 740 |
| QY | 171 PMAAPEVIOSLPVSEFTCDTYSYGVVLWEMLTREVPFKGLEGLQVAWLVEKNERLTIPSS    | 230 |
| Db | 741 EWAAPPEVLRRDEPSNEKSDVYSFGVILWELATLLOQPCWCNLPQAQVAAVGFKGKRLDIPRD | 800 |
| QY | 231 CPRSFAELHQCEADAKKRPSPFKOIIISILESMS                              | 264 |
| Db | 801 VNPKLASLIVACWADPEPWKRPSFSSIMETLKPMT                             | 834 |

```

RESULT 8
US-09-840-704-5
; Sequence 5, Application US/09840704
; Patent No. US20020122801A1
; GENERAL INFORMATION:
; APPLICANT: Dedhar, Shoukat
; APPLICANT: Hannigan, Greg
; TITLE OF INVENTION: Intergrin-Linked Kinase and its Uses
; FILE REFERENCE: KIN-2CON
; CURRENT APPLICATION NUMBER: US/09/840,704
; CURRENT FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 09/566,906
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: US08/752,345
; PRIOR FILING DATE: 1996-11-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 263
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: Other
; LOCATION: (1)...(263)
US-09-840-704-5

```

```

Query Match      20.5%; Score 487; DB 10; Length 263;
Best Local Similarity 41.0%; Pred. No. 1.1e-28;
Matches 109; Conservative 43; Mismatches 98; Indels 16; Gaps 5;
QY      11 IKFDLQFFENCGGSFGSVYRAKMTISQDKEAVAKKLKIE-----KEAEILSVL 60
      | : * | | | | | | | | | | | | | | | | | | | | : | | : |

```

[illegible]

```

RESULT 9
US-10-158-895-4
; Sequence 4, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-4

```

| Query Match           | 20.3%;       | Score 481;   | DB 9;           | Length 579; |
|-----------------------|--------------|--|-----------------|-------------|
| Best Local Similarity | 30.4%;       | Pred. No. 7.9e-28;   |                 |             |
| Matches 132;          | Conservative | 81;  | Mismatches 151; | Indels 70;  |
|                       |              |  |                 | Gaps 18;    |
| QY                    | 7            | SEVQIKFDLDLQFFENCGGSGSVYRAKWSIQDKEVAVKLLIKIEKAE-----IL         | 57              |             |
| DB                    | 27           | NFEEDIDYKEIEVEEAVVGRGAFGVYCKAKW--RAKDVAIK---QIESESERKAFIVELROL | 81              |             |
| QY                    | 58           | SVLSHRNIIQFYGVILEPPNYGIVTEYASLGLSYDYINSNREEM---DMDHIMTATDV     | 114             |             |
| DB                    | 82           | SRVNHPNIVKLYGACLN--VCLIMEVAEGSGSLYNVLHG--AEPDPYYTAHAAMSWCLOQ   | 137             |             |
| QY                    | 115          | AKGMWYILMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRPHNHTTHMSLVGTFPW   | 172             |             |
| DB                    | 138          | SGGVAYLHSMQPKALIHRLDKRPNUILLVAGTVLKICDFGTACDIQTHMTNNK--GSAAM   | 195             |             |
| QY                    | 173          | MAPEVIQSLPVSETCDTYSYGVVLEMWLTVREVPFKGLEG--LOVAVLVEKNERLTIPTSS  | 230             |             |
| DB                    | 196          | MAPEVREGSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN   | 254             |             |
| QY                    | 231          | CPRSFALLHQCEWADAKKRPSEFKQIISIL-----ESMSNDT-                    | 267             |             |
| DB                    | 255          | LPKPIESLMTRCWSKDPSPQRPSEMEIYKIMTHLMRYFPGADEPLQYPCQYSDGQSN SAT  | 314             |             |
| QY                    | 268          | ---SLPDKCNSFLHNKAEMRCE-IEATLERLKLKERDLSFKEQELKERERRLKMWEQKLT   | 323             |             |
| DB                    | 315          | STGSEMDIASTNTSNKSDTNNMEQVPATNDTIKLESKLLKNQAKQOSESGRSL---GAS    | 371             |             |

QY 324 EQSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQITATSNGEHGMNPSLOAMLMGF 383  
DB 372 HGSSVESLPTS-----EGKMSADMSEIARIAATT-GNGQPRRRSIQDLTVTGT 421  
QY 384 --GDIFSMNKAGAV 395  
DB 422 EPGQVSSRSSSPSV 435

## RESULT 10

US-10-158-895-15  
; Sequence 15, Application US/10158895  
; Patent No. US20020155624A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-158-895-15

Query Match 20.3%; Score 481; DB 9; Length 590;  
Best Local Similarity 30.4%; Pred. No. 8.1e-28;  
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 7 SFVQIKFDLQFFENGCGSGSVYRAKMSQDKEVAVKLLKIEKAE-----IL 57  
DB 27 NFEIDYKEIEVEEVGRGAFGVCAKAK--RAKDAIK--QIESESEKAFIVELOL 81  
QY 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEM--DMDHMTATDV 114  
DB 82 SRVNHPIVLYGACLP--VCLVMEYAEAGSLYNVLAHG--AEPLPYTAHAHMSWCLQC 137  
115 AKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFHNHTHMSLVGTFPW 172  
138 SGGVAYLHSMQPKALIHRLDKPNNLLVAGTGLKICDFGACDIQTHMTNNK--GSAAW 195  
QY 173 MAPEVIQSLPVSETCDTYSYGVVLMEMLTREVPFKGLEG--LQVAVLVEKNERLTIPSS 230  
DB 196 MAPEVFEGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIXN 254  
QY 231 CPSPFAELLHQCEWADAKKRPSPKQIISIL-----ESMSNDT- 267  
DB 255 LKPIESLMTRCWSKDPQSPQSMEEIVKIMTHLMRYFPGADEPLQYPCQYSDGQSN SAT 314  
QY 268 ---SLPDKCNFLHNKAWEKCE-IEATLERLKKLERDLSFKEQELKERERRLKMEQKLT 323  
DB 315 STGSFMDIASTNTSNKSDPTNMEQVPATNDTIKRLSKLLKNQAKQOSESGRSL--GAS 371  
QY 324 EQSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQITATSNGEHGMNPSLOAMLMGF 383  
DB 372 HGSSVESLPTS-----EGKMSADMSEIARIAATT-GNGQPRRRSIQDLTVTGT 421  
QY 384 --GDIFSMNKAGAV 395  
DB 422 EPGQVSSRSSSPSV 435

RESULT 11

US-09-764-868-799  
; Sequence 799, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 799  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (4)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-868-799

Query Match 17.9%; Score 424; DB 9; Length 92;  
Best Local Similarity 98.8%; Pred. No. 1.4e-24;  
Matches 80; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 53 EAILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTAT 112  
DB 12 QAILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTAT 71  
QY 113 DVAKGMHYLHMEAPVKVIHRD 133  
DB 72 DVAKGMHYLHMEAPVKVIHRD 92

RESULT 12  
US-09-947-199-8  
; Sequence 8, Application US/09947199  
; Patent No. US20020127684A1  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jayaseelan  
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCE: NMI-068CP2  
; CURRENT APPLICATION NUMBER: US/09/947,199  
; CURRENT FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: 60/111,938  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/291,839  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: 09/458,457  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 835  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-947-199-8

Query Match 16.5%; Score 392.5; DB 10; Length 835;  
Best Local Similarity 29.6%; Pred. No. 4.6e-21;  
Matches 115; Conservative 61; Mismatches 151; Indels 61; Gaps 12;

QY 11 IKFDLQFFENGCGSGSVYRAKMSQDKEVAVK-----LTKIEKAEILSV 59  
DB 458 LQISEIFHEIIGSGSFQYVYKGR--CRNKIVAIKRYRANTYCSKSDVDMFCREVSILCQ 515  
QY 60 LSHRNIIQFYGVILEPPN-YGIVTEYASLSGLYDINSNRSEEMDMHMTATDVAKGM 118  
DB 516 LNHPCVQVFGACLDPSQFAIVTQYISGSLFSLHLEQK-RILDQSKLIIADVAKGM 574  
QY 119 HYLH-MEAPVKVIHRDLKSRNVVIAADGVLKICDFGASRFHNHTHMSLV---GTFPMA 174



Db 575 EYLHSLTOP--IHRDLNSHNILLYEDGHAVNADGESRFLQSLDEDNNTKOPGNLRWMA 632  
QY 175 PEVI-QSLPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAMLVKERNERTIPSSCPR 233  
Db 633 PEVFTQCTRYTIKADVFSYSLCLMELLTGEIPFAHLKPAADAMAYHHIRPIGYSIPK 692  
QY 234 SFAELLHQWEADAKRPSFKQIISILE-----SMSNDTSLPDKNS--FLHN 279  
Db 693 PISLLIRGNWACPEGRPEFSEYVSKLEBCLCNVELMSPASSNSSGSLSPSSSDCLLSR 752  
QY 280 KAEMRCEIEATLERLKLRLDSFKEQLKERERRLKMWEOQLTEQSNTPLLPLAARMS 339  
Db 753 GGPGRSHVAALRSR-----ELEYALNARSYAGWSQSVGTHSNPGLSLEMMNRST 802  
QY 340 EES-----YFESKTEESN 352  
Db 803 QYSTVDKYGVSDDPMSLTHLHSRQDDSN 830

ULT 13  
09-771-161A-255  
; Sequence 255, Application US/09771161A  
; Patent No. US20020110811A1  
; GENERAL INFORMATION:  
; APPLICANT: LEVINE, et al.  
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
; FILE REFERENCE: 802620-2005.1  
; CURRENT APPLICATION NUMBER: US/09/771,161A  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/724,676  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 136776  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 135619  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 273  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 255  
; LENGTH: 1036  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-771-161A-255

Query Match 16.2%; Score 385; DB 10; Length 1036;  
Best Local Similarity 33.1%; Pred. No. 2.1e-20;  
Matches 104; Conservative 52; Mismatches 116; Indels 42; Gaps 13;

Db 15 DLQFFENCGGSGFSVYRAKM-ISQDEYAVK-KLKIE-----KEAIIISVL 60  
QY 61 SHRNIQFYGVILEPPNYGIVTEYASIGSLYDINSNRSEMDMHIMTATDVAKGMHY 120  
Db 185 DHRNLIRLYGVLTTP-MKMTTELAPLGSLDLRLKRGHFL-LGTLTRYAVQVAEGMGY 242  
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASR-FHNHTTMSLVG----TFPWWAP 175  
Db 243 LESK--RFIHRDLAARNLLATRDVLKIGDFGLMRALPQNDHYVMQEHKRVPAWCAP 299  
QY 176 EVIQSLPVSETCDTYSYGVLWEMLT-REVPFKGLEGLQVAMLVKERNERTIPSSCPRS 234  
Db 300 ESLKTRTFSHASDTWTFGVTLMEMFTYQGEPIWIGLNGSQILHKIDKEGERLPRPDCPOD 359  
QY 235 FAELLHQWEADAKRPSFKQIIS-ILSMSNDTSL-----PDKCN-----SFLHNK 280  
Db 360 IYNVWVQCWAHKPEDRPTFVALRDFLLEAQPTDMRALQDFEERPKLHIQMNVDITVIEGR 419  
QY 281 AE--WRCEIEATL 291  
Db 420 AENYWMRGQNTRTL 433

RESULT 14  
US-09-771-161A-256  
; Sequence 256, Application US/09771161A  
; Patent No. US20020110811A1  
; GENERAL INFORMATION:  
; APPLICANT: LEVINE, et al.  
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
; FILE REFERENCE: 802620-2005.1  
; CURRENT APPLICATION NUMBER: US/09/771,161A  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/724,676  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 136776  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 135619  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 273  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 256  
; LENGTH: 1036  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-771-161A-256

Query Match 16.2%; Score 385; DB 10; Length 1036;  
Best Local Similarity 33.1%; Pred. No. 2.1e-20;  
Matches 104; Conservative 52; Mismatches 116; Indels 42; Gaps 13;

QY 15 DLQFFENCGGSGFSVYRAKM-ISQDEYAVK-KLKIE-----KEAIIISVL 60  
Db 125 DLRLLEKLGDSFGVVRGEMDAPSGKTVSAVKCLKPDVLSQPEAMDFIREVNAMHSL 184  
QY 61 SHRNIQFYGVILEPPNYGIVTEYASIGSLYDINSNRSEMDMHIMTATDVAKGMHY 120  
Db 185 DHRNLIRLYGVLTTP-MKMTTELAPLGSLDLRLKRGHFL-LGTLTRYAVQVAEGMGY 242  
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASR-FHNHTTMSLVG----TFPWWAP 175  
Db 243 LESK--RFIHRDLAARNLLATRDVLKIGDFGLMRALPQNDHYVMQEHKRVPAWCAP 299  
QY 176 EVIQSLPVSETCDTYSYGVLWEMLT-REVPFKGLEGLQVAMLVKERNERTIPSSCPRS 234  
Db 300 ESLKTRTFSHASDTWTFGVTLMEMFTYQGEPIWIGLNGSQILHKIDKEGERLPRPDCPOD 359  
QY 235 FAELLHQWEADAKRPSFKQIIS-ILSMSNDTSL-----PDKCN-----SFLHNK 280  
Db 360 IYNVWVQCWAHKPEDRPTFVALRDFLLEAQPTDMRALQDFEERPKLHIQMNVDITVIEGR 419  
QY 281 AE--WRCEIEATL 291  
Db 420 AENYWMRGQNTRTL 433

RESULT 15  
US-08-987-689A-32  
; Sequence 32, Application US/08987689A  
; Patent No. US20020048782A1  
; GENERAL INFORMATION:  
; APPLICANT: Sima Lev  
; APPLICANT: Joseph Schlessinger  
; TITLE OF INVENTION: PYK2 RELATED PRODUCTS AND METHODS  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0

```
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/987,689A
FILING DATE: December 9, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,824
FILING DATE: December 11, 1996
APPLICATION NUMBER: 08/460,626
FILING DATE: June 2, 1995
APPLICATION NUMBER: 08/357,642
FILING DATE: December 15, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Wardburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 230/110
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 251 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-987-689A-32
```

```
Query Match 16.1%; Score 382.5; DB 8; Length 251;
Best Local Similarity 33.9%; Pred. No. 5.8e-21;
Matches 83; Conservative 51; Mismatches 96; Indels 15; Gaps 7;

QY 23 GGSFSGSVYRAKWSQDKEVAVKCL---LKIE---KEAELSVLSHRNIIQFYGVILEP 75
DB 7 GGGQYGEVYEGWKKYSLTVAVKTLKEDTWEVEEFLKEAAVMKEIKHPNLVQLLGCTRE 66

QY 76 PNYGI VTEYASLSLYDIYINSRSEEMDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLK 135
DB 67 PPFYITTEFTYGNLLDYLRECNREQEVAVVLLYMATQISSAMEYLEKK--NFIHRDLA 123

QY 136 SRNVVIADGVLCICDFGASRFHNHTHMSLVGT-FP--WMAPEVIOQLPVSETCDITYSY 192
DB 124 ARNCLVGENHIVKVDGFLSRMTGDTYTAHAGAKFPKWTAPESLAYNKFSIKSDVWAF 183

QY 193 GAWLWEMLTREV-PFKGLEGLQVAMLVKEKNERLTIPSSCPRSFAELLHQCWEDAKKRP 251
DB 184 GVLWELIATYGMSPYPGIDRSQV-YELLEKDYRMKRPEGCEKVEYELMRACWQWNPSDRP 242

252 SFKQI 256
243 SFAEI 247
```

Search completed: May 1, 2003, 20:52:12  
Job time : 35.1911 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:33:12 ; Search time 30.4603 Seconds  
(without alignments)  
439.505 Million cell updates/sec

Title: US-09-757-982-5  
Perfect score: 2375  
Sequence: 1 MSLGASFVQIKFDLQFFE.....GDDDDDDGEEEDNDMDNSE 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues  
al number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep: \*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep: \*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep: \*  
6: /cgn2\_6/ptodata/1/iaa/backfillseq1.pep: \*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID               | Description        |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1          | 2375  | 100.0       | 455    | 3 US-09-221-235-5   | Sequence 5, Appli  |
| 2          | 2375  | 100.0       | 455    | 3 US-09-221-928-5   | Sequence 5, Appli  |
| 3          | 2375  | 100.0       | 455    | 4 US-09-221-527-5   | Sequence 5, Appli  |
| 4          | 2375  | 100.0       | 455    | 4 US-09-221-236-5   | Sequence 5, Appli  |
| 5          | 2375  | 100.0       | 455    | 4 US-09-221-416-5   | Sequence 5, Appli  |
| 6          | 2375  | 100.0       | 455    | 4 US-09-221-245-5   | Sequence 5, Appli  |
| 7          | 2375  | 100.0       | 455    | 4 US-09-163-115-5   | Sequence 5, Appli  |
| 8          | 2375  | 100.0       | 455    | 4 US-09-221-528-5   | Sequence 5, Appli  |
| 9          | 2375  | 100.0       | 455    | 4 US-09-593-553-5   | Sequence 5, Appli  |
| 10         | 2375  | 100.0       | 455    | 4 US-09-221-237-5   | Sequence 5, Appli  |
| 11         | 565.5 | 23.8        | 668    | 1 US-08-205-018-2   | Sequence 2, Appli  |
| 12         | 565.5 | 23.8        | 859    | 1 US-08-395-580-2   | Sequence 2, Appli  |
| 13         | 565.5 | 23.8        | 859    | 5 PCT-US95-02792-2  | Sequence 2, Appli  |
| 14         | 488   | 20.5        | 821    | 1 US-07-928-464-2   | Sequence 2, Appli  |
| 15         | 488   | 20.5        | 821    | 1 US-08-003-311B-2  | Sequence 2, Appli  |
| 16         | 488   | 20.5        | 821    | 1 US-08-261-432-2   | Sequence 2, Appli  |
| 17         | 488   | 20.5        | 821    | 5 PCT-US93-07347-2  | Sequence 2, Appli  |
| 18         | 487   | 20.5        | 263    | 3 US-09-035-706-5   | Sequence 5, Appli  |
| 19         | 487   | 20.5        | 263    | 3 US-08-955-841-5   | Sequence 5, Appli  |
| 20         | 487   | 20.5        | 263    | 4 US-09-390-425-5   | Sequence 5, Appli  |
| 21         | 487   | 20.5        | 263    | 4 US-09-566-906-5   | Sequence 5, Appli  |
| 22         | 481   | 20.3        | 579    | 4 US-09-529-279-4   | Sequence 4, Appli  |
| 23         | 481   | 20.3        | 590    | 4 US-09-529-279-15  | Sequence 15, Appli |
| 24         | 401.5 | 16.9        | 1584   | 4 US-09-457-040B-27 | Sequence 27, Appli |
| 25         | 397.5 | 16.7        | 275    | 2 US-08-701-191A-36 | Sequence 36, Appli |
| 26         | 393.5 | 16.6        | 261    | 2 US-07-857-224B-59 | Sequence 59, Appli |
| 27         | 391.5 | 16.5        | 261    | 2 US-07-857-224B-60 | Sequence 60, Appli |

|    |       |      |     |                     |                   |
|----|-------|------|-----|---------------------|-------------------|
| 28 | 379.5 | 16.0 | 505 | 1 US-08-222-616-20  | Sequence 20, Appl |
| 29 | 379.5 | 16.0 | 505 | 4 US-08-446-648-20  | Sequence 20, Appl |
| 30 | 379.5 | 16.0 | 505 | 5 PCT-US95-04228-20 | Sequence 20, Appl |
| 31 | 377.5 | 15.9 | 506 | 4 US-08-426-509A-6  | Sequence 6, Appli |
| 32 | 377.5 | 15.9 | 511 | 5 PCT-US95-05008-6  | Sequence 6, Appli |
| 33 | 377   | 15.9 | 835 | 4 US-09-291-839-2   | Sequence 2, Appli |
| 34 | 376.5 | 15.9 | 537 | 4 US-08-426-509A-11 | Sequence 11, Appl |
| 35 | 376.5 | 15.9 | 537 | 5 PCT-US95-05008-11 | Sequence 11, Appl |
| 36 | 376   | 15.8 | 283 | 2 US-08-701-191A-41 | Sequence 41, Appl |
| 37 | 374.5 | 15.8 | 259 | 2 US-07-857-224B-52 | Sequence 52, Appl |
| 38 | 365.5 | 15.4 | 536 | 4 US-08-426-509A-12 | Sequence 12, Appl |
| 39 | 365.5 | 15.4 | 536 | 5 PCT-US95-05008-12 | Sequence 12, Appl |
| 40 | 364   | 15.3 | 820 | 1 US-08-166-717D-6  | Sequence 4, Appli |
| 41 | 358.5 | 15.1 | 675 | 4 US-08-426-509A-4  | Sequence 6, Appli |
| 42 | 358.5 | 15.1 | 675 | 5 PCT-US95-05008-4  | Sequence 4, Appli |
| 43 | 358   | 15.1 | 729 | 1 US-07-640-029-3   | Sequence 3, Appli |
| 44 | 358   | 15.1 | 731 | 1 US-07-921-807B-5  | Sequence 5, Appli |
| 45 | 358   | 15.1 | 731 | 1 US-08-441-944A-5  | Sequence 5, Appli |

ALIGNMENTS

|  |     |  |     |  |  |  |  |  |  |
|--|-----|--|-----|--|--|--|--|--|--|
| RESULT 1   |     |  |     |  |  |  |  |  |  |
| US-09-221-235-5  |     |  |     |  |  |  |  |  |  |
| ; Sequence 5, Application US/09221235  |     |  |     |  |  |  |  |  |  |
| ; Patent No. 6043040   |     |  |     |  |  |  |  |  |  |
| ; GENERAL INFORMATION:   |     |  |     |  |  |  |  |  |  |
| ; APPLICANT: Acton, Susan  |     |  |     |  |  |  |  |  |  |
| ; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR |     |  |     |  |  |  |  |  |  |
| ; FILE REFERENCE: MNI-050  |     |  |     |  |  |  |  |  |  |
| ; CURRENT APPLICATION NUMBER: US/09/221, 235                                 |     |  |     |  |  |  |  |  |  |
| ; CURRENT FILING DATE: 1998-12-28  |     |  |     |  |  |  |  |  |  |
| ; EARLIER APPLICATION NUMBER: 09/163,115                                     |     |  |     |  |  |  |  |  |  |
| ; EARLIER FILING DATE:   |     |  |     |  |  |  |  |  |  |
| ; NUMBER OF SEQ ID NOS: 15   |     |  |     |  |  |  |  |  |  |
| ; SOFTWARE: Patentin Ver. 2.0  |     |  |     |  |  |  |  |  |  |
| ; SEQ ID NO 5  |     |  |     |  |  |  |  |  |  |
| ; LENGTH: 455  |     |  |     |  |  |  |  |  |  |
| ; TYPE: PRT  |     |  |     |  |  |  |  |  |  |
| ; ORGANISM: Homo sapiens   |     |  |     |  |  |  |  |  |  |
| ; US-09-221-235-5  |     |  |     |  |  |  |  |  |  |
| Query Match  |     |  |     |  |  |  |  |  |  |
| Best Local Similarity 100.0%; Score 2375; DB 3; Length 455;                  |     |  |     |  |  |  |  |  |  |
| Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;                 |     |  |     |  |  |  |  |  |  |
| QY   | 1   | MSLGASFVQIKFDLQFFENCGGSGSVYRAKWSIQDKEVAVKLLKIEKAELSVL        | 60  |  |  |  |  |  |  |
| DB   | 1   | MSLGASFVQIKFDLQFFENCGGSGSVYRAKWSIQDKEVAVKLLKIEKAELSVL        | 60  |  |  |  |  |  |  |
| QY   | 61  | SHRNIIQFYGVILEPPNYGIVTEYASLSGLVDYINSNRSEEMDMHMTWATDVAKGMHY   | 120 |  |  |  |  |  |  |
| DB   | 61  | SHRNIIQFYGVILEPPNYGIVTEYASLSGLVDYINSNRSEEMDMHMTWATDVAKGMHY   | 120 |  |  |  |  |  |  |
| QY   | 121 | LHMEAPVKVIHRDLKSRNVVIAADGVLLKICDFGASRFHNHTTHMSLVGTPWMAPEVIQS | 180 |  |  |  |  |  |  |
| DB   | 121 | LHMEAPVKVIHRDLKSRNVVIAADGVLLKICDFGASRFHNHTTHMSLVGTPWMAPEVIQS | 180 |  |  |  |  |  |  |
| QY   | 181 | LPVSETCDTVSYGVVLMEMLTREVPFKGLEGLQVAVLVEKNERLTTPSSCPRSFAELH   | 240 |  |  |  |  |  |  |
| DB   | 181 | LPVSETCDTVSYGVVLMEMLTREVPFKGLEGLQVAVLVEKNERLTTPSSCPRSFAELH   | 240 |  |  |  |  |  |  |
| QY   | 241 | QCWEADAKKRPSPKQIISILESMSNDTSLPDKNSFLHNKAEMWRCETATLERLKLKLRD  | 300 |  |  |  |  |  |  |
| DB   | 241 | QCWEADAKKRPSPKQIISILESMSNDTSLPDKNSFLHNKAEMWRCETATLERLKLKLRD  | 300 |  |  |  |  |  |  |
| QY   | 301 | LSFKQELKERERRLLKMEQKLTQSNTPLLPLAARMSEESYFESKTEBSNSAEMSCQI    | 360 |  |  |  |  |  |  |
| DB   | 301 | LSFKQELKERERRLLKMEQKLTQSNTPLLPLAARMSEESYFESKTEBSNSAEMSCQI    | 360 |  |  |  |  |  |  |
| QY   | 361 | TATSNGEHGHNPSLQAMLMGFGDIFSNNKAGAVMHSQMQINQAKONSSKTSKRGRK     | 420 |  |  |  |  |  |  |



Db 361 TATSNGEHGMNPSLQAMLMGFGDIFSMNKAGAVMHSQMIMQAKONSSKTSKRKG 420  
QY 421 KVNMALGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455  
Db 421 KVNMALGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 2  
US-09-221-928-5  
; Sequence 5, Application US/09221928  
; Patent No. 6121030  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,928  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5

; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-928-5

Query Match 100.0%; Score 2375; DB 3; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.3e-192;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSQDKEVAVKKLKEKEAETLSVL 60  
Db 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSQDKEVAVKKLKEKEAETLSVL 60  
QY 61 SHRNIQFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
Db 61 SHRNIQFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEQLQVAVLVEKNERLTI PSSCPRSFAELLH 240  
Db 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEQLQVAVLVEKNERLTI PSSCPRSFAELLH 240  
QY 241 QCWEADAKKRPSFKQIISLESMSNDTSLPDKCNSFLHNKAEMWRCETLRLKLERD 300  
Db 241 QCWEADAKKRPSFKQIISLESMSNDTSLPDKCNSFLHNKAEMWRCETLRLKLERD 300  
QY 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCOI 360  
Db 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCOI 360  
QY 361 TATSNGEHGMNPSLQAMLMGFGDIFSMNKAGAVMHSQMIMQAKONSSKTSKRKG 420  
Db 361 TATSNGEHGMNPSLQAMLMGFGDIFSMNKAGAVMHSQMIMQAKONSSKTSKRKG 420  
QY 421 KVNMALGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455  
Db 421 KVNMALGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 3  
US-09-221-527-5  
; Sequence 5, Application US/09221527  
; Patent No. 6146832  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,527

; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-527-5

Query Match 100.0%; Score 2375; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.3e-192;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSQDKEVAVKKLKEKEAETLSVL 60  
Db 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSQDKEVAVKKLKEKEAETLSVL 60  
QY 61 SHRNIQFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
Db 61 SHRNIQFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEQLQVAVLVEKNERLTI PSSCPRSFAELLH 240  
Db 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEQLQVAVLVEKNERLTI PSSCPRSFAELLH 240  
QY 241 QCWEADAKKRPSFKQIISLESMSNDTSLPDKCNSFLHNKAEMWRCETLRLKLERD 300  
Db 241 QCWEADAKKRPSFKQIISLESMSNDTSLPDKCNSFLHNKAEMWRCETLRLKLERD 300  
QY 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCOI 360  
Db 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCOI 360  
QY 361 TATSNGEHGMNPSLQAMLMGFGDIFSMNKAGAVMHSQMIMQAKONSSKTSKRKG 420  
Db 361 TATSNGEHGMNPSLQAMLMGFGDIFSMNKAGAVMHSQMIMQAKONSSKTSKRKG 420  
QY 421 KVNMALGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455  
Db 421 KVNMALGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 4  
US-09-221-236-5  
; Sequence 5, Application US/09221236  
; Patent No. 6146841  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,236  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-236-5

Query Match 100.0%; Score 2375; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.3e-192;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSQDKEVAVKKLKEKEAETLSVL 60

Db 1 MSSLGASFVQIKFDLQFFENCGGSGFSVYRAKWIISQDKEVAVKCLKLKEKEAETLSVL 60  
Qy 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMDHIMTWATDVAKGMY 120  
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMDHIMTWATDVAKGMY 120  
Qy 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180  
Qy 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVLVVEKNERLTISSSCPSPFAELH 240  
Db 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVLVVEKNERLTISSSCPSPFAELH 240  
Qy 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIEATLERLKLKLERD 300  
Db 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIEATLERLKLKLERD 300  
Qy 301 LSFKEQELKERERRLKWEQKLTQSNTPLLPLAARMSSESYFESKTEESNSAEMSCQI 360  
Db 301 LSFKEQELKERERRLKWEQKLTQSNTPLLPLAARMSSESYFESKTEESNSAEMSCQI 360  
Qy 361 TATSNGEHGMNPSLOAMLMGFCDIFSMNKAGAVMHSGMOINMOAKONSSKTSKRGRK 420  
Db 361 TATSNGEHGMNPSLOAMLMGFCDIFSMNKAGAVMHSGMOINMOAKONSSKTSKRGRK 420  
Qy 421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455  
Db 421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

RESULT 5

US-09-221-416-5  
; Sequence 5, Application US/09221416  
; Patent No. 6153417  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,416  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-221-416-5

Query Match 100.0%; Score 2375; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.3e-192;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSSLGASFVQIKFDLQFFENCGGSGFSVYRAKWIISQDKEVAVKCLKLKEKEAETLSVL 60  
Db 1 MSSLGASFVQIKFDLQFFENCGGSGFSVYRAKWIISQDKEVAVKCLKLKEKEAETLSVL 60  
Qy 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMDHIMTWATDVAKGMY 120  
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMDHIMTWATDVAKGMY 120  
Qy 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180  
Qy 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVLVVEKNERLTISSSCPSPFAELH 240  
Db 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVLVVEKNERLTISSSCPSPFAELH 240  
Qy 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIEATLERLKLKLERD 300  
Db 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIEATLERLKLKLERD 300

Db 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIEATLERLKLKLERD 300  
Qy 301 LSFKEQELKERERRLKWEQKLTQSNTPLLPLAARMSSESYFESKTEESNSAEMSCQI 360  
Db 301 LSFKEQELKERERRLKWEQKLTQSNTPLLPLAARMSSESYFESKTEESNSAEMSCQI 360  
Qy 361 TATSNGEHGMNPSLOAMLMGFCDIFSMNKAGAVMHSGMOINMOAKONSSKTSKRGRK 420  
Db 361 TATSNGEHGMNPSLOAMLMGFCDIFSMNKAGAVMHSGMOINMOAKONSSKTSKRGRK 420  
Qy 421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455  
Db 421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

RESULT 6

US-09-221-245-5  
; Sequence 5, Application US/09221245  
; Patent No. 6180358  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,245  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: US 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-221-245-5

Query Match 100.0%; Score 2375; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.3e-192;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSSLGASFVQIKFDLQFFENCGGSGFSVYRAKWIISQDKEVAVKCLKLKEKEAETLSVL 60  
Db 1 MSSLGASFVQIKFDLQFFENCGGSGFSVYRAKWIISQDKEVAVKCLKLKEKEAETLSVL 60  
Qy 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMDHIMTWATDVAKGMY 120  
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEMDMDHIMTWATDVAKGMY 120  
Qy 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180  
Qy 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVLVVEKNERLTISSSCPSPFAELH 240  
Db 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVLVVEKNERLTISSSCPSPFAELH 240  
Qy 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIEATLERLKLKLERD 300  
Db 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIEATLERLKLKLERD 300  
Qy 301 LSFKEQELKERERRLKWEQKLTQSNTPLLPLAARMSSESYFESKTEESNSAEMSCQI 360  
Db 301 LSFKEQELKERERRLKWEQKLTQSNTPLLPLAARMSSESYFESKTEESNSAEMSCQI 360  
Qy 361 TATSNGEHGMNPSLOAMLMGFCDIFSMNKAGAVMHSGMOINMOAKONSSKTSKRGRK 420  
Db 361 TATSNGEHGMNPSLOAMLMGFCDIFSMNKAGAVMHSGMOINMOAKONSSKTSKRGRK 420  
Qy 421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455  
Db 421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

```
RESULT 7
US-09-163-115-5
; Sequence 5, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-163-115-5

Query Match          100.0%; Score 2375; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.3e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWSQDKEVAVKKLKEAEILSVL 60
Db 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWSQDKEVAVKKLKEAEILSVL 60

QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEMDMDHMTWATDVAKGMHY 120
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEMDMDHMTWATDVAKGMHY 120

QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTMSLVGTFPWWAPEVIO 180
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTMSLVGTFPWWAPEVIO 180

QY 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPRSFAELLH 240
Db 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPRSFAELLH 240

QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIATLERLKKLERD 300
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIATLERLKKLERD 300

QY 301 LSFKEQLKERERRLKMEQKLTQOSNTPLLLPLARMSEESYFESKTEESNSAEMSCQI 360
Db 301 LSFKEQLKERERRLKMEQKLTQOSNTPLLLPLARMSEESYFESKTEESNSAEMSCQI 360

QY 361 TATSNGEHGMNPSLOAMLMGFDFISMNKAAGAVMHSQMGINMOAKONSSKTSKRGRK 420
Db 361 TATSNGEHGMNPSLOAMLMGFDFISMNKAAGAVMHSQMGINMOAKONSSKTSKRGRK 420

QY 421 KVNMAIGFSDFDLSEGDDDDDDGEEEDNDMDNSE 455
Db 421 KVNMAIGFSDFDLSEGDDDDDDGEEEDNDMDNSE 455

RESULT 8
US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-09-221-528-5

Query Match          100.0%; Score 2375; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.3e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWSQDKEVAVKKLKEAEILSVL 60
Db 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWSQDKEVAVKKLKEAEILSVL 60

QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEMDMDHMTWATDVAKGMHY 120
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEMDMDHMTWATDVAKGMHY 120

QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTMSLVGTFPWWAPEVIO 180
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTMSLVGTFPWWAPEVIO 180

QY 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPRSFAELLH 240
Db 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPRSFAELLH 240

QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIATLERLKKLERD 300
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAEMRCEIATLERLKKLERD 300

QY 301 LSFKEQLKERERRLKMEQKLTQOSNTPLLLPLARMSEESYFESKTEESNSAEMSCQI 360
Db 301 LSFKEQLKERERRLKMEQKLTQOSNTPLLLPLARMSEESYFESKTEESNSAEMSCQI 360

QY 361 TATSNGEHGMNPSLOAMLMGFDFISMNKAAGAVMHSQMGINMOAKONSSKTSKRGRK 420
Db 361 TATSNGEHGMNPSLOAMLMGFDFISMNKAAGAVMHSQMGINMOAKONSSKTSKRGRK 420

QY 421 KVNMAIGFSDFDLSEGDDDDDDGEEEDNDMDNSE 455
Db 421 KVNMAIGFSDFDLSEGDDDDDDGEEEDNDMDNSE 455

RESULT 9
US-09-593-553-5
; Sequence 5, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-593-553-5

Query Match          100.0%; Score 2375; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.3e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWSQDKEVAVKKLKEAEILSVL 60
Db 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWSQDKEVAVKKLKEAEILSVL 60

QY 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEMDMDHMTWATDVAKGMHY 120
Db 61 SHRNIIOFYGVILEPPNYGIVTEYASLSGLYDINSNRSEMDMDHMTWATDVAKGMHY 120

QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTMSLVGTFPWWAPEVIO 180
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTMSLVGTFPWWAPEVIO 180
```



Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLKICDFGASRFHNHTTHMSLVGTFPWWAPEVIQS 180  
QY 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAMLVYEKNERLTISSCPRSFAELLH 240  
Db 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAMLVYEKNERLTISSCPRSFAELLH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATLERLKLKERD 300  
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATLERLKLKERD 300  
QY 301 LSFKEQELKERERRLKMWEOKLTQSNTPLLPLAARMSSESYFESKTEESNSAEMSCQI 360  
Db 301 LSFKEQELKERERRLKMWEOKLTQSNTPLLPLAARMSSESYFESKTEESNSAEMSCQI 360  
QY 361 TATSNGEHGMNPSLOAMLMGFGDIFSMNKAGAVHSGMQINMOAKONSSKTSKRGRK 420  
Db 361 TATSNGEHGMNPSLOAMLMGFGDIFSMNKAGAVHSGMQINMOAKONSSKTSKRGRK 420  
QY 421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455  
Db 421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

## RESULT 10

US-09-221-237-5  
; Sequence 5, Application US/09221237  
; Patent No. 6214597  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,237  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OR SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-237-5

Query Match 100.0%; Score 2375; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.3e-192;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKMSQDKEVAVKLLIKEAEILSVL 60  
1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKMSQDKEVAVKLLIKEAEILSVL 60  
QY 61 SHRNIIOFYGVILEPENVGIVTEYASLSGLYDINSNRSEEMDMHIMTWATVAKGMHY 120  
Db 61 SHRNIIOFYGVILEPENVGIVTEYASLSGLYDINSNRSEEMDMHIMTWATVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLKICDFGASRFHNHTTHMSLVGTFPWWAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLKICDFGASRFHNHTTHMSLVGTFPWWAPEVIQS 180  
QY 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAMLVYEKNERLTISSCPRSFAELLH 240  
Db 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAMLVYEKNERLTISSCPRSFAELLH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATLERLKLKERD 300  
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATLERLKLKERD 300  
QY 301 LSFKEQELKERERRLKMWEOKLTQSNTPLLPLAARMSSESYFESKTEESNSAEMSCQI 360  
Db 301 LSFKEQELKERERRLKMWEOKLTQSNTPLLPLAARMSSESYFESKTEESNSAEMSCQI 360  
QY 361 TATSNGEHGMNPSLOAMLMGFGDIFSMNKAGAVHSGMQINMOAKONSSKTSKRGRK 420  
Db 361 TATSNGEHGMNPSLOAMLMGFGDIFSMNKAGAVHSGMQINMOAKONSSKTSKRGRK 420

Db 361 TATSNGEHGMNPSLOAMLMGFGDIFSMNKAGAVHSGMQINMOAKONSSKTSKRGRK 420  
QY 421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455  
Db 421 KVMALGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

## RESULT 11

US-08-205-018-2  
; Sequence 2, Application US/08205018  
; Patent No. 5554523  
; GENERAL INFORMATION:  
; APPLICANT: Reddy, Usharani R.  
; APPLICANT: Pleasure, David  
; TITLE OF INVENTION: No. 5554523e1 Protein Kinase, Nucleic Acid  
; TITLE OF INVENTION: Sequences Encoding the Same and Methods Related  
; TITLE OF INVENTION: Thereto  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and  
; ADDRESSEE: No. 5554523ris  
; STREET: One Liberty Place - 46th floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/205,018  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gaumont, Rebecca R.  
; REGISTRATION NUMBER: 35,152  
; REFERENCE/DOCKET NUMBER: CH-0488  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 668 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-205-018-2

Query Match 23.8%; Score 565.5; DB 1; Length 668;  
Best Local Similarity 37.8%; Pred. No. 1.5e-39;  
Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;

QY 10 QIKFDDLOFFENC GGSGSVYRAKMSQDKEVAVKLLIKEAEI--LSVLSHRNIIO 67  
Db 119 EVPFEEILDLOWVSGAQAQAVFLGRF--HGEVAVKVRDL-KETDIKHLRKLKHPNIIT 175  
QY 68 FYGVILEPENVGIVTEYASLSGLYDINSNRSEEMDMHIMTWATVAKGMHYLHMEAPV 127  
Db 176 FKGVCTQAPCYCLIMEFCAQGLYEVLRAGRPTPSL--LVDWSMGIAGMNYLHLH-- 230  
QY 128 KVIHRDLKSRNVVIAADGVLKICDFGASR-FHNHTTHMSLVGTFPWWAPEVIQS LPVSET 186  
Db 231 KTIHRDLKSPNMLTYDDVVKISDFGTSKELSDKSTKMSFAGTVAMMAPEVIRNBPVSEK 290  
QY 187 CDTYSYGVLWEMLTREVPFKGLEGLQVAMLVYEKNERLTISSCPRSFAELLHQCWEAD 246  
Db 291 VDIWSFGVWLWELLTGEIPIKDVDSAIHWGVSNSLHLVPSSCPDGFKILLRQCWNSK 350  
QY 247 AKKRPSFKQIISILESMSNDT-SLPDKCNFLHNKAERCEIATLERLK-----KLE 298  
Db 351 PRNRPSFRQILHLHDIASADVLTPOE--TYFKSOAEWRREVKLHFEEKISBGTCLHRL 408





OY 299 RDLSEKOE-----LKERERLKMEOKLTQSNTPLLPLAARMSEESYFESKTEE 350  
Db 409 EELVMRRRELRHALDIREH----YERKLERANN--LYMELNALMLQ---LELKERE 456

## RESULT 14

US-07-928-464-2  
Sequence 2, Application US/07928464  
Patent No. 5367065  
GENERAL INFORMATION:  
APPLICANT: Ecker, Joseph R.  
APPLICANT: Kieber, Joseph J.  
TITLE OF INVENTION: Constitutive Triple Response Gene and  
TITLE OF INVENTION: Mutations  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz and  
ADDRESSEE: No. 5367065r1s  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/928,464  
FILING DATE: 19920810  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Miller, Suzanne E.  
REGISTRATION NUMBER: 32,279  
REFERENCE/DOCKET NUMBER: UPN-1086  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 821 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-928-464-2

Query Match 20.5%; Score 488; DB 1; Length 821;  
Best Local Similarity 40.7%; Pred. No. 6.9e-33;  
Matches 109; Conservative 44; Mismatches 99; Indels 16; Gaps 5;

OY 9 VOIKFDLQFFENCGGSGFSGVYRAKWSQDKEVAVKLLKIE-----KEAEILS 58  
Db 544 MDIPWCDLNIEKIGAGSFGTVHRAEWHGSD--VAVKILMEQDFHAERVNEFLREVAIMK 601  
OY 59 VLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYN--SNRSEEMDMHMTATDVAKG 117  
Db 602 RLHRPNIVLFMGAVTQPNLSIVTEYLSRGLYRLHLHKSAREQDLDERRRRLSMAYDVAKG 661  
OY 118 MHYLMHAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHT--THMSLVGTFFWMAAP 175  
Db 662 MNYLHNRP-PIVHRDLKSPNLLVDKYYTVKVCDFGLSRLKASTFLSSKSAAGTPPEWMAAP 720  
OY 176 EVIQLPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPRSF 235  
Db 721 EYLRDEPSNEKSDVYSFGVILMELATLQQPWGNLPAQVAAVGFCKCKLEIPRNLPQV 780  
OY 236 AELLHQCWEADAKRPSFKQIISILESM 263  
Db 781 AALIEGCTNEPWKRPSFATIMDLRPL 808

RESULT 15  
US-08-003-311B-2

Sequence 2, Application US/08003311B  
Patent No. 544166  
GENERAL INFORMATION:

APPLICANT: Ecker, Joseph R.  
APPLICANT: Kieber, Joseph J.  
TITLE OF INVENTION: Constitutive Triple Response Gene  
TITLE OF INVENTION: and Mutations  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and  
ADDRESSEE: No. 544166r1s  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/003,311B  
FILING DATE: January 12, 1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/928,464  
FILING DATE: August 10, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lori Y. Beardsell  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: UPN-1108  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 821 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-003-311B-2

Query Match 20.5%; Score 488; DB 1; Length 821;  
Best Local Similarity 40.7%; Pred. No. 6.9e-33;  
Matches 109; Conservative 44; Mismatches 99; Indels 16; Gaps 5;

OY 9 VOIKFDLQFFENCGGSGFSGVYRAKWSQDKEVAVKLLKIE-----KEAEILS 58  
Db 544 MDIPWCDLNIEKIGAGSFGTVHRAEWHGSD--VAVKILMEQDFHAERVNEFLREVAIMK 601  
OY 59 VLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYN--SNRSEEMDMHMTATDVAKG 117  
Db 602 RLHRPNIVLFMGAVTQPNLSIVTEYLSRGLYRLHLHKSAREQDLDERRRRLSMAYDVAKG 661  
OY 118 MHYLMHAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHT--THMSLVGTFFWMAAP 175  
Db 662 MNYLHNRP-PIVHRDLKSPNLLVDKYYTVKVCDFGLSRLKASTFLSSKSAAGTPPEWMAAP 720  
OY 176 EVIQLPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPRSF 235  
Db 721 EYLRDEPSNEKSDVYSFGVILMELATLQQPWGNLPAQVAAVGFCKCKLEIPRNLPQV 780  
OY 236 AELLHQCWEADAKRPSFKQIISILESM 263  
Db 781 AALIEGCTNEPWKRPSFATIMDLRPL 808

Search completed: May 1, 2003, 20:39:29  
Job time : 37.4603 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:31:42 ; Search time 14.8131 Seconds  
(without alignments)  
1602.986 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_31\_277  
Perfect score: 1300  
Sequence: 1 YRAKWISQDKEVAVKLLKI.....SILESMSNDTSLPDKCNSFL 247

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 511.5 | 39.3        | 406    | 2     | T52626      |
| 2          | 510.5 | 39.3        | 394    | 2     | TJ00229     |
| 3          | 495.5 | 38.1        | 328    | 2     | T16747      |
| 4          | 490.5 | 37.7        | 954    | 1     | S68178      |
| 5          | 490   | 37.7        | 668    | 2     | JC2363      |
| 6          | 490   | 37.7        | 888    | 2     | A55318      |
| 7          | 488   | 37.5        | 888    | 2     | JC5399      |
| 8          | 482.5 | 37.1        | 847    | 1     | A53800      |
| 9          | 480   | 36.9        | 886    | 2     | T48544      |
| 10         | 475.5 | 36.6        | 1030   | 2     | F96763      |
| 11         | 460.5 | 35.4        | 462    | 2     | S29851      |
| 12         | 457.5 | 35.2        | 390    | 2     | T01451      |
| 13         | 453   | 34.8        | 982    | 2     | T06576      |
| 14         | 448   | 34.5        | 412    | 2     | T10671      |
| 15         | 442   | 34.0        | 821    | 2     | T48400      |
| 16         | 436.5 | 33.6        | 1015   | 2     | T00726      |
| 17         | 434   | 33.4        | 848    | 2     | B87950      |
| 18         | 434   | 33.4        | 855    | 2     | T20082      |
| 19         | 431.5 | 33.2        | 553    | 2     | T04683      |
| 20         | 431   | 33.2        | 829    | 2     | T07406      |
| 21         | 427   | 32.8        | 475    | 2     | T12955      |
| 22         | 420.5 | 32.3        | 546    | 2     | D84555      |
| 23         | 418   | 32.2        | 407    | 2     | G84635      |
| 24         | 417   | 32.1        | 736    | 2     | T05137      |
| 25         | 403.5 | 31.0        | 988    | 2     | F86316      |
| 26         | 402   | 30.9        | 963    | 2     | T09911      |
| 27         | 396.5 | 30.5        | 545    | 2     | T05675      |
| 28         | 395   | 30.4        | 567    | 2     | JC5957      |
| 29         | 395   | 30.4        | 579    | 2     | JC5955      |

|    |       |      |      |   |        |                    |
|----|-------|------|------|---|--------|--------------------|
| 30 | 395   | 30.4 | 606  | 2 | JC5956 | transforming growt |
| 31 | 381.5 | 29.3 | 738  | 2 | F96701 | hypothetical prote |
| 32 | 369.5 | 28.4 | 1130 | 1 | TVHUA  | protein-tyrosine k |
| 33 | 369   | 28.4 | 1257 | 2 | T00486 | serine/threonine-s |
| 34 | 368.5 | 28.3 | 981  | 1 | FOMVGM | gag-ab1 polypeptei |
| 35 | 368.5 | 28.3 | 1123 | 2 | A39962 | kinase-related tra |
| 36 | 367   | 28.2 | 364  | 2 | G71410 | probable protein k |
| 37 | 365.5 | 28.1 | 1146 | 2 | B35962 | protein-tyrosine k |
| 38 | 365.5 | 28.1 | 1182 | 2 | A35962 | protein-tyrosine k |
| 39 | 364.5 | 28.0 | 1520 | 1 | TVFEFA | protein-tyrosine k |
| 40 | 362.5 | 27.9 | 357  | 2 | C84856 | probable protein k |
| 41 | 362.5 | 27.9 | 391  | 2 | T48115 | protein kinase ATM |
| 42 | 362   | 27.8 | 842  | 2 | T32258 | hypothetical prote |
| 43 | 358.5 | 27.6 | 1147 | 2 | F86297 | hypothetical prote |
| 44 | 355.5 | 27.3 | 1584 | 2 | T18276 | protein-tyrosine k |
| 45 | 354.5 | 27.3 | 1171 | 2 | T12956 | hypothetical prote |

ALIGNMENTS

RESULT 1

T52626

probable mitogen-activated protein kinase MAP3K delta-1 (imported) - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 08-Dec-2000

C/Accession: T52626

R/Jouanvic, S.; Hamal, A.; Lepince, A.S.; Tregear, J.W.; Krets, M.; Henry, Y.

Gene 229, 171-81, 1999

A/Title: Characterisation of novel plant genes encoding MEKK/STE11 and RAF-related prot

A/Reference number: Z24447; MUID:99196996; PMID:10095117

A/Accession: T52626

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-406 <JOU>

A/Cross-references: EMBL.Y14199; NID:g2253009; PIDN:CAA74591.1; PID:g2253010

A/Experimental source: cultivar Columbia

C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

Query Match 39.3%; Score 511.5; DB 2; Length 406;  
Best Local Similarity 40.5%; Pred. No. 1.8e-23;  
Matches 104; Conservative 51; Mismatches 85; Indels 17; Gaps 6;

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | YRAKWISQDKEVAVKLL-----KIEKEAELSVLSHRNIIQYGVILEPPNYGI           | 50  |
| DB | 150 | YRAEW--NGTEVAVKFLDQDFSGDALTFKSEIEIMRLRHPNVVLFMGAVTRP NFSI      | 207 |
| QY | 51  | VTEYASLGSLDYINSNRSEEMDMHTWATDVAKGMHYLHMEAPVKYIHRDLKSRNV        | 110 |
| DB | 208 | LTEFLPRGSLYRLH-RPNHQLDEKRRMVALDVAKGMNYLHTSHT-VVHRDLKSPNLL      | 265 |
| QY | 111 | IAADGVLKICDFGASRFHNHT--THMSLVGTFPMAPEVIGSLVSETCDTYSYGVLMWE     | 168 |
| DB | 266 | VDKNMVVVKVCDFGLSRMKHTTYSKSTAGTPEMAPVLRNRPANEKCDVYSFGVILWE      | 325 |
| QY | 169 | MLTRVFPKGLGLEQVAMLVVEKNERLTIIPSSCPSPFAELHQCWEADAKKRPSEKQIIS    | 228 |
| DB | 326 | LATSRVPWKGLNPMQVGVAGVGFQNRRLIEIPDDIDLTVAQIIRCEWQTEPHLRPSFTQLMQ | 385 |
| QY | 229 | ILESMSNDTSLPDKCNS 245  |     |
| DB | 386 | SLKRLQG-LNISNRANT 401  |     |

RESULT 2

TJ00229

mixed-lineage protein kinase 1 - human

C/Species: Homo sapiens (man)

C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 17-Mar-1999

C/Accession: S32467; JU0229

R/Dorow, D.S.; Devetieux, L.; Dietzsch, E.; de Kretser, T.

Eur. J. Biochem. 213, 701-710, 1993

A/Title: Identification of a new family of human epithelial protein kinases containing





|    |     |   |     |
|----|-----|---|-----|
| Db | 227 | NILILEAIENHNILADTVLKITDFGLAREWHKTKMSAAGTYAMMAPEVIRLISLFSKSSDV | 286 |
| Qy | 160 | YSYGVVLEWMLTREVPFKGLEQVAVLVEKNERLTIIPSSCPRSFALLHQWEADAKK      | 219 |
| Db | 287 | WSFGVLMELLTGEVPPYREIDALAVAGVAMNKLTLPIPTCPEPFARLLEECDPDPHG     | 346 |
| Qy | 220 | RPSFKOITSILESMSNDTSLPDKCSNF                                   | 246 |
| Db | 347 | RPDEGSILKRLEVIEOSALFQMPLESF                                   | 373 |

## RESULT 5

protein kinase (EC 2.7.1.37) ZPK - human  
N/Alternate names: leucine-zipper protein kinase  
C/Species: Homo sapiens (man)  
C/Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Feb-2001  
C/Accession: J02363  
C/Author: Reddy, U.R.; Pleasure, D.  
C/Journal: J. Biol. Chem. 267: 613-620, 1994  
C/Title: Cloning of a novel putative protein kinase having a leucine zipper domain from  
A/Reference number: J02363; MUID:94311945; PMID:8037767

A/Accession: J02363  
A/Molecule type: mRNA  
A/Residues: 1-668 <RED>  
A/Cross-references: EMBL:U07358  
A/Experimental source: brain  
A/Note: the nucleotide sequence for this amino acid sequence is inconsistent with that for the codon ACC for residue 661 as pro, the codon GAAACACCTCTCCA for residues 664-668 as A  
C/Comment: This protein belongs to the family of non-receptor kinase.  
C/Genetics:

A;Gene: GDB:ZPK  
A;Cross-references: GDB:383963; OMIM:600447

C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
 A/Map position: 12q13-12q13  
 C/Keywords: ATP; leucine zipper; nucleotide binding; P-loop; phosphotransferase  
 F/123-371/Domain: protein kinase homology <KIN>  
 F/131-139/Region: protein kinase ATP-binding motif  
 F/443-471/Region: leucine zipper motif  
 F/538-545/Region: nucleotide-binding motif A (P-loop)  
 F/152/Active site: Lys #status predicted

|                       |                  |                    |            |             |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match           | 37.7%            | Score 490;         | DB 2;      | Length 668; |
| Best Local Similarity | 44.3%            | Pred. No. 5.5e-22; |            |             |
| Matches 105;          | Conservative 43; | Mismatches 79;     | Indels 10; | Gaps 6;     |

```

10 KEVAVKCLKLIEKEAEI--LSVLSHRNIIOFYGVILEPNNYGIYTEYASLSGLSDYDINSN 67
   :|||||:  : |||: | | | | | | | | | | | | | | | | | | | | | | | | | | |
147 EYAVVKVRDL-KEFDIKHLRKLKHPNIITFGVCTOAPCYCIIMEFCAOGOLYEYLRAG 205

```

```
QY      68 RSEEMDMHIMTWATDVAKGMHYLHNEAPVKIHRDLKSRNVIAADGVLCIDFGASR- 126
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      206 RPVTPTSL--LVDWSMGIAAGMNYLHIH---KIHRDLKSPNNITVYDNIARKSDNECSYE 360
```

QY 127 FHNHTNSLVGTFPWMAPEVIQSLPVS ETCDTYSYGVVIM EMLTREVPFKGLEGLQVAM 186  
: : | | | | | | | | : : | | | | | | : : | | : |  
Db 261 LSDKSTKISFAGTVAMMAPEVINRPEVSEKVDIWSFGVIMELLTGEIRPYKDVSSAIIW 320

QY 187 LVVEKNERLTIPSSCPRSFAELLHQCEADAKKRPSPFOIISILESMSNDT-SLPDK 242  
| : : ||||| | | : : : ||| : : | : | : | :  
Db 321 GVGSGNSLHPVPSSCPDGFKILLRQCWNSKPRNRPSPFOILLHLDIASADVSTPQE 377

## RESULT 6

serine/threonine protein kinase (EC 2.7.1.-) DLK - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 24-Sep-1999  
C/Accession: A55318  
R/Holzman, L.B.; Merritt, S.E.; Fan, G.  
J. Biol. Chem. 269, 30808-30817, 1994  
A/Title: Identification, molecular cloning, and characterization of dual leucine zipper

9.  
A:Reference number: A55318; MUID:95074107; PMID:7983011

A;Accession: A55318  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-888 <HO

A;Cross-references: GB:U14636; NID:g602677; PIDN:AAA57280.1; PID:g602678  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C/Keywords: ATP; leucine zipper; phosphotransferase  
F;156-404/Domain: protein kinase homology <KIN>  
F;164-172/Region: protein kinase ATP-binding motif

|             |        |            |       |             |
|-------------|--------|------------|-------|-------------|
| Query Match | 37.7%; | Score 490; | DB 2; | Length 888; |
|-------------|--------|------------|-------|-------------|

best local similarity 44.3%; Pred. No. 7.2e-22;  
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

```
OY      10 KEVAVKKLLKIEKEAEI--LSVLSHRNIIQFYGVILEPPNVGIVTEYASLGSGLDYINSN    67  
       :|::||:::||::||::||::||::||::||::||::||::||::||::||::||::||  
Db     180 EEVAVKKVVDL-KETDIKHRLRKHKHNIITFKGVCTQAPEYCILMEFCAQGQLYEVIIRAG   238
```

```
QY      68 RSEEMDMHIMTWATDVAKGMHYLMEAPVKVIHRDLKSRNVIIADGVLKICDEGASR-   126
         | : :: |||::||| :|:|||||::||| :|:||||| :
Db     239 RPVTPEL--LVDSWMSGIAGMNYHLH---KIIRDLKSPNMLITYDDVVKISDFGTSKE 293
```

QY 127 FHNHTHNSLVGTFPMMAPAEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLEQVAV 186  
: : | | | | | | | : : | | | | | : : :  
Db 294 LSDKSTKNSFAGTVAMMAPEVIRNEPVSEKVDIWSFGVLMELLTGEIPYKVDSSAIIW 353

```
Qy 187 LVEKNERLTTPSSCPRSFAELLHQCEADAKKRPSFKQIISILESMSNDT-SLPDK 242
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 GVGSNLSHLPEVSSCPDGFKILLRQCNMCKPRNRPSPFRQILLHLDIASAVLSTPQE 410
```

## RESULT 7

UC5399  
dual leucine zipper kinase (EC 2.7.--.) - rat

C:species: katusus norvegicus (Norway rat)  
C:Date: 07-Jul-1997 #sequence\_18-Jul-1997 #text\_change 28-May-1999  
C:Accession: J05399

R. Matsui, N.; Sarkar, G.; Shuto, T.; Marrs, J.; Bronk, J.T.; Mizuno, K.; Bolander, M.E. Biochem. Biophys. Res. Commun. 229, 571-576, 1996

A; Reference number: JC5399; MUID:97127443; PMID:89549339

A:Accession: UC5399  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA

A:Residues: 1-888 <MAT>  
 C:Comment: This enzyme is involved in regulating cell function in the musculoskeletal system.  
 C:Genetics: chr1  
 A:Gene: rnr

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol,  
C;Keywords: ATP; phosphotransferase  
F;54-75, 89-98/Region: glycine-rich  
F;156-454/Domain: kinase catalytic #status predicted <CAT>

F:156-404/Domain: protein kinase  
F:164-172/Region: protein kinase ATP-binding motif

|                   |   |
|-------------------|---|
| F;421-449/Region: | leucine zipper motif                          |
| F;472-500/Region: | leucine zipper motif                          |
| F;557-888/Region: | glycine-serine-proline rich #status predicted |

|                           |        |                    |            |             |
|---------------------------|--------|--------------------|------------|-------------|
| Query Match               | 37.5%; | Score 488;         | DB 2;      | Length 888; |
| Best Local Similarity     | 43.9%; | Pred. No. 9.4e-22; |            |             |
| Matches 104; Conservative | 44;    | Mismatches 79;     | Indels 10; | Gaps 6.     |

Qy 1

Db 180 EEVAVKVRDL-KETDIKILKLIKHPNIIITFKGVCTQAPCYILMEFCAQGQLYEVLRAQ 238

Qy 68 RSEEMDMDHINTWATDVAKGMHYLLHMEAPVKVIHRDLKSRNVVIAADGVLKICDFGASR- 126

127 FHNHTTMSLVGTFEPMABEVIQSLPVSETCDTYSYGVLMEMLTREVPFKGLEGOVAM 186

Db 294 LSDKSTKMSFAGTAVAMADPEVIRNEPVSEKVDIMSGVLLWELLTGEIPYKDVDSALITW 353  
OY 187 LVEKNERLTIPSSCPSPRSFAELHQCWEADAKRSPKQIISILBSMSNDT-SLPDK 242  
Db 354 GVGSNLSHLPPVSSCPDGFKILLRQCWNSKPRNPSFRQILLHLDIASADVLTSPQE 410

RESULT 8

A53800  
mixed-lineage protein kinase (EC 2.7.1.-) 3 - human  
N/Alternate names: protein kinase PTK1; protein kinase SPRK  
C/Species: Homo sapiens (man)  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C/Accession: A53800; 158395  
R/Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.  
J. Biol. Chem. 269, 15092-15100, 1994  
A/Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont  
A/Reference number: A53800; MUID:94253068; PMID:8195146  
Accession: A53800  
Status: preliminary  
Molecule type: mRNA  
A/Residues: 1-847 <GAL>  
A/Cross-references: GB:U07747; NID:g464027; PIDN:AAA19647.1; PID:g464028  
R/Ing, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lassam, N.J.  
Oncogene 9, 1745-1750, 1994  
A/Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domai  
A/Reference number: 158395; MUID:94239754; PMID:8183572  
A/Accession: 158395  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-847 <RES>  
A/Cross-references: GB:L32976; NID:g488295; PIDN:AAA59859.1; PID:g488296  
C/Genetics:  
A/Gene: GDB:MLK3; PTK1; SPRK  
A/Cross-references: GDB:134755; OMIM:600050  
A/Map position: 11q13.1-11q13.3  
C/Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology  
C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k  
F/48-100/Domain: SH3 homology <SH32>  
F/115-383/Domain: protein kinase homology <KIN>  
F/123-131/Region: protein kinase ATP-binding motif  
F/403-424/Region: leucine zipper motif  
F/438-459/Region: leucine zipper motif  
F/468-482/Region: basic

Query Match 37.1%; Score 482.5; DB 1; Length 847;  
Best local similarity 41.1%; Pred. No. 1.9e-21;  
Matches 104; Conservative 44; Mismatches 80; Indels 25; Gaps 6;  
OY 1 YRAKWISQ-----DKEAVKLLKIEKEAELSVLSHRNIIQFYGVILEPPNYG 49  
Db 132 YRGSWRGELVAVKARQDPDEDISV-TAESVQEARLFAMLAHPNIIALKAVCLEEPNLC 190  
OY 50 IVEYASLSGLDYINSNRSEEMDMDHIM-TWATDVAKGMHYLHMEAPVKVIHRDLKSRN 108  
Db 191 LVMEYAAAGPLSRALAGRVP----PHLVNMAVQIARGMHYLHCEALVPVIRDLKSN 246  
OY 109 VV-----IAADGV---LKICDFGASRFHNHTTHMSLVGTFPWMAPEVIGSLPVSETCDTY 160  
Db 247 ILLIQPIESDDMEHKTLDTFGLAREMHKTQMSAAGTYAWMAPEVIKASTFSKSDVW 306  
OY 161 SYGVVLMEMLTREVPFKGLEGLQVAMLVEKNERLTIPSSCPSPRSFAELHQCWEADAKR 220  
Db 307 SFGVLLWELLTGEVPPYRGIDCLAVAYGVAVNKLTLPIPTCPEPFAQLMADCAQDPHRR 366  
OY 221 PSFKQIISILESM 233  
Db 367 PDFASILQQLBAL 379

RESULT 9  
T48544

MAP3K delta-1 protein kinase - Arabidopsis thaliana  
N/Alternate names: protein F14F18.20  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C/Accession: T48544  
R/Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,  
submitted to the Protein Sequence Database, April 2000  
A/Reference number: Z2490  
A/Accession: T48544  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-886 <BEV>  
A/Cross-references: EMBL:AL163812  
A/Experimental source: cultivar Columbia; BAC clone F14F18  
C/Genetics:  
A/Map position: 5  
A/Introns: 148/3; 180/3; 327/3; 362/2; 568/3; 658/1; 671/3; 694/3; 728/2; 754/3; 810/3;  
A/Note: F14F18.20

Query Match 36.9%; Score 480; DB 2; Length 886;  
Best local similarity 42.9%; Pred. No. 2.8e-21;  
Matches 97; Conservative 41; Mismatches 72; Indels 16; Gaps 5;

OY 1 YRAKWISQDKEVAVKKLL-----KIEKEAELSVLSHRNIIQFYGVILEPPNYGI 50  
Db 664 YRAEW--NGTEVAVKKFLDQDFSGDALTFKSEIEMRLRHPNVLLFMGAVTRPPNFSI 721  
OY 51 VTEYASLSGLDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNV 110  
Db 722 LTFELPRGSLYRLIH-RPNHQLDEKRMALDVAKGMHYLHSHPT-VVHRDLKSPNLL 779  
OY 111 IAADGVLKICDFGASRFHNHT--THMSLVGTFPWMAPEVIGSLPVSETCDTYSYGVVLM 168  
Db 780 VDKNWVVKVCDPFGISRKHHTYLSKSTAGTPEWMAPEVLRNEPANKECDVYSFGVILWE 839  
OY 169 MLTREVPFKGLEGLQVAMLVEKNERLTIPSSCPSPRSFAELHQCWE 214  
Db 840 LATSRLVPWKGLNPMQVGVAVGFQNRRLIIPDDIDLTVAQIIRECWQ 885

RESULT 10

F96763  
hypotheoretical protein F25P22.8 (imported) - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: F96763  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K,  
ansen, N.F.; Hughes, B.; Huiztar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: F96763  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-1030 <STO>  
A/Cross-references: GB:AE005173; NID:g6692730; PIDN:AAF24836.1; GSPDB:GN00141  
C/Genetics:  
A/Gene: F25P22.8  
A/Map position: 1

Query Match 36.6%; Score 475.5; DB 2; Length 1030;  
Best local similarity 38.3%; Pred. No. 5.9e-21;  
Matches 102; Conservative 48; Mismatches 85; Indels 31; Gaps 7;

OY 1 YRAKWISQDKEVAVKKLL-----KIEKEAELSVLSHRNIIQFYGVILEPPNYGI 50  
Db 763 YRGDW--HGTAVAVKKFLDQDLTGEALIEFRSEVRIMKLRHPNIVLFMGAVTRPPNLSI 820



[illegible]

## RESULT 11

protein kinase 6 (EC 2.7.1.-) - soybean  
C/Species: Glycine max (soybean)  
C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 24-Sep-1999  
C/Accession: S29851, S27760  
R/Feng, X.H.; Zhao, Y.; Bottino, P.J.; Kung, S.  
Biochim. Biophys. Acta 1172, 200-204, 1993  
A/Title: Cloning and characterization of a novel member of protein kinase family from sc  
A/Reference number: S29851, MUID:93176812, PMID:8439562  
A/Accession: S29851  
A/Molecule type: mRNA  
A/Residues: 1-462 <FEN>  
A/Cross-references: EMBL:M67449; NID:g170046; PIDN:AAA34002.1; PID:g170047  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C/Keywords: ATP; phosphotransferase  
F154-419/Domain: protein kinase homology <KIN>

|                       |                  |                    |            |             |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match           | 35.4%            | Score 460.5;       | DB 2;      | Length 462; |
| Best Local Similarity | 37.0%            | Pred. No. 2.1e-20; |            |             |
| Matches 98;           | Conservative 53; | Mismatches 81;     | Indels 33; | Gaps 7;     |

```

QY 6 ISQDEVAVKLL-----KIEK----EABILSVLSHRNIOFYGVILEPPNYG 49
Db 174 VYKDEAVAVKIMVPEDDDGNGLASRLKQFIREVTLLSRLHHQNVIKFSACRKPVC 233
QY 50 IVTEYASLSGLDYINSNRSEEMDMHIMTATDVAKGMHYLHNEAPVKVIHRDLKSRNV 109
234 IITEYLAEGSLRAYLHKLHEQITSLQKLIAPALDIARGMEYIHSOG--VIHRDLKPENI 290
110 VIAADVLKICDFGASRFHNHTHMSLV---GTFPMAPAEVIOSLPVSETCDTYSYGVV 165
Db 291 LINEDNHLKIADFGIA---CEERASCDLLADDPGTYRMMAPEMIKRKSYGKKVDVYSFGLI 347
QY 166 LMEMLTREVPFKGLELOVAVLVEKNERLTISSCPRSPAEILLHQCEWADAKKRPSFKQ 225
Db 348 LMEMLTGTIPYEDMNPIDQAFVAVVNKNSRPTIPSNCPAMRALIEQCWSLQDPDKRPEFWQ 407
QY 226 IISILE---SMSNDTSL---PDKC 243
Db 408 VKILIEQFESSLASDGTISLVENPC 432

```

## RESULT 12

protein kinase homolog F2401.13 - *Arabidopsis thaliana*  
C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001  
C/Accession: T01451  
R/Shim, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Cor-  
eologs, A.; Ecker, J.R.  
submitted to the EMBL Data Library, January 1998  
A/Description: Genomic sequence for *Arabidopsis thaliana* BAC F2401.  
A/Reference number: Z14211  
A/Accession: T01451

A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-390 <SH1>  
A;Cross-references: EMBL:AC003113; NID:g2689438; PID:g2781357; GSPDB:GN00059; ATSP:F240  
C;Genetics:  
A;Gene: ATSP:F240L.13  
A;Map position: 1  
A;Introns: 149/3; 301/3  
C;Superfamily: kinase-related transforming protein; protein kinase homolog

|                       |        |                    |                |             |
|-----------------------|--------|--------------------|----------------|-------------|
| Query Match           | 35.2%  | Score 457.5;       | DB 2;          | Length 390; |
| Best Local Similarity | 37.2%; | Pred. No. 2.7e-20; |                |             |
| Matches               | 96;    | Conservative 49;   | Mismatches 86; | Indels 27.  |

[illegible]

## RESULT 13

probable protein kinase TC1R2 - tomato  
C/Species: Lycopersicon esculentum (tomato)  
C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 08-Oct-1999  
C/Accession: T06576  
R/Hackett, R.M.  
submitted to the EMBL Data Library, March 1998  
A/Reference number: Z15770  
A/Accession: T06576  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-982 <HAC>  
A/Cross-references: EMBL:AJ005077; NID:e1296722; PIDN:CAA06334.1; PID:e1296723  
A/Experimental source: cultivar Ailsa craig  
C/Genetics:  
A/Gene: TC1R2

|                           |       |                    |            |             |
|---------------------------|-------|--------------------|------------|-------------|
| Query Match               | 34.8% | Score 453;         | DB 2;      | Length 982; |
| Best Local Similarity     | 38.5% | Pred. No. 1.2e-19; |            |             |
| Matches 101; Conservative | 45;   | Mismatches 96;     | Indels 30; | Cuts 2      |

```

QY      1 YRAKWSIQDEKVAAYKKLL-----KIEKEAELSVLSHRNITQFYGVILEPPNYGI 50
Db      716 YHADW--NGTEVAAYKKFLDQDFSGAALAEFRKEVRIMRRLRHPNVAVREMGAITRPPHLSI 777
QY      51 VTEYASLGLSYDYINSNRSEBMDMDHIMTATDVAKGMHYLHMEAPYKVIHRDLKSRNV 110
Db      774 ITEFLPRGSLYRIIHRPHF-QIDERQKIKMALDVAKGMDCLHTSNPT-IVHRDLKSPNLT 831
QY      111 IAADGLKICDFGASRFHNHT--THMSLVGTFFPMABEVIQSLPVSETCDTYSYGVWLWE 168
Db      832 VDTDMNVKVCDFGLSRLLKHTNFTLSKSKSTAGTPEMMAPEVLKRNESNEKCDIYSFGVILWE 891
QY      169 MLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPRSFAELHOCWEADAKKRPSFKOITS 228
Db      892 LATRLPWSGNNPMQVVGAVGFQNRKLEIPKELDPIVARIITWECWOTDPNLRPSPFAQLTV 951

```



```
QY 229 ILESMSNDTSLP---DKCNSFL 247
      | : : | | |
Db 952 ALTPLOQ-LVIPAYVDQLSRL 972
```

## RESULT 14

T10671  
protein kinase homolog F6E21.90 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 19-May-2000  
C;Accession: T10671  
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft  
submitted to the Protein Sequence Database, June 1999  
A;Reference number: Z16533  
A;Accession: T10671  
A;Molecule type: DNA  
A;Residues: 1-412 <BEV>  
A;Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21.90  
A;Experimental source: cultivar Columbia; BAC clone F6E21  
Genetics:  
Gene: ATSP:F6E21.90  
Map position: 4  
A;Introns: 300/2  
C;Superfamily: Kinase-related transforming protein; protein kinase homology  
F;135-392/Domain: protein kinase homology <KIN>

## RESULT 15

T48400  
serine/threonine-protein kinase ctrl - Arabidopsis thaliana  
N;Alternate names: protein F17C15.150  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 17-Nov-2000  
C;Accession: T48400; A45178  
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;  
submitted to the Protein Sequence Database, March 2000  
A;Reference number: Z24492  
A;Accession: T48400  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-821 <BEV>  
A;Cross-references: EMBL:AL162506  
A;Experimental source: cultivar Columbia; BAC clone F17C15  
R;Kieber, J.U.; Rothenberg, M.; Roman, G.; Feldmann, K.A.; Ecker, J.R.  
Cell 72, 427-441, 1993  
A;Title: ~~GR1~~, a negative regulator of the ethylene response pathway in Arabidopsis, enc  
A;Reference number: A45178; MUID:93161417; PMID:8431946  
A;Accession: A45178

A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: DNA  
A/Residues: 1-468, 470-821 <KIE>  
A/Note: sequence extracted from NCBI backbone (NCBIP:124878)

|                       |                  |                    |            |             |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match           | 34.0%;           | Score 442;         | DB 2;      | Length 821; |
| Best Local Similarity | 40.2%;           | Pred. No. 4.5e-19; |            |             |
| Matches 99;           | Conservative 41; | Mismatches 90;     | Indels 16; | Gaps 5;     |

Search completed: May 1, 2003, 20:38:39  
Job time : 16.8131 secs

GenCore version 5.1.4 ps\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:13:36 ; Search time 12.4017 Seconds

(without alignments)  
826.070 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_31\_277  
Perfect score: 1300  
Sequence: 1 YRAKMWISQDKEVAVKKLLKI.....SILESMSNDTSLPDKCNFL 247

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5 ;

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID           | Description        |
|------------|-------|-------------|--------|--------------|--------------------|
| 1          | 510.5 | 39.3        | 394    | 1 M3K9_HUMAN | P80192 homo sapien |
| 2          | 490.5 | 37.7        | 954    | 1 M3KA_HUMAN | Q02779 homo sapien |
| 3          | 490   | 37.7        | 859    | 1 M3KC_HUMAN | Q12852 homo sapien |
| 4          | 490   | 37.7        | 888    | 1 M3KC_MOUSE | Q60700 mus musculu |
| 5          | 488   | 37.5        | 888    | 1 M3KC_RAT   | Q63796 rattus norv |
| 6          | 442   | 34.0        | 821    | 1 CTR1_ARATH | Q05609 arabidopsis |
| 7          | 395   | 30.4        | 579    | 1 M3K7_MOUSE | Q62073 mus musculu |
| 8          | 395   | 30.4        | 606    | 1 M3K7_HUMAN | Q43318 homo sapien |
| 9          | 369.5 | 28.4        | 1130   | 1 ABL1_HUMAN | P00519 homo sapien |
| 10         | 368.5 | 28.3        | 746    | 1 ABL_MLVAB  | P00521 abelson mur |
| 11         | 368.5 | 28.3        | 1123   | 1 ABL1_MOUSE | P00520 mus musculu |
| 12         | 365.5 | 28.1        | 1182   | 1 ABL2_HUMAN | P42684 homo sapien |
| 13         | 364.5 | 28.0        | 1520   | 1 ABL_DROME  | P00522 drosophila  |
| 14         | 359.5 | 27.3        | 1584   | 1 KYK1_DICDI | P18160 dictyosteli |
| 15         | 351.5 | 27.0        | 439    | 1 ABL_FSVHY  | P10447 feline sarc |
| 16         | 349.5 | 26.9        | 393    | 1 M3K7_MOUSE | P83104 drosophila  |
| 17         | 345   | 26.5        | 536    | 1 FYN_XIPH   | P27446 xiphophorus |
| 18         | 343   | 26.4        | 505    | 1 FRK_HUMAN  | P42685 homo sapien |
| 19         | 337.5 | 26.0        | 536    | 1 FYN_XENLA  | P13406 xenopus lae |
| 20         | 335.5 | 25.8        | 536    | 1 FYN_HUMAN  | P06241 homo sapien |
| 21         | 335   | 25.8        | 819    | 1 FGRI_CHICK | P21804 gallus gall |
| 22         | 335   | 25.8        | 822    | 1 FGRI_HUMAN | P11362 homo sapien |
| 23         | 335   | 25.8        | 822    | 1 FGRI_MOUSE | P16092 mus musculu |
| 24         | 334.5 | 25.7        | 533    | 1 FYN_MOUSE  | P39688 mus musculu |
| 25         | 332.5 | 25.6        | 587    | 1 SRC_AVIS2  | P15054 mus musculu |
| 26         | 332   | 25.5        | 517    | 1 FGR_MOUSE  | P14234 avian sarco |
| 27         | 330   | 25.4        | 410    | 1 KYK2_DICDI | P18161 dictyosteli |
| 28         | 330   | 25.4        | 822    | 1 FGRI_RAT   | Q04589 rattus norv |
| 29         | 329.5 | 25.3        | 1196   | 1 ABL1_CAEBL | P03949 caenorhabdi |
| 30         | 329   | 25.3        | 806    | 1 CEK2_CHICK | P18460 gallus gall |
| 31         | 328.5 | 25.3        | 450    | 1 CSK_HUMAN  | P41240 homo sapien |
| 32         | 328.5 | 25.3        | 532    | 1 SRC_CHICK  | P00523 gallus gall |
| 33         | 327.5 | 25.2        | 531    | 1 SRC1_XENLA | P13115 xenopus lae |

|    |       |      |     |              |                    |
|----|-------|------|-----|--------------|--------------------|
| 34 | 327.5 | 25.2 | 531 | 1 SRC2_XENLA | P13116 xenopus lae |
| 35 | 327   | 25.2 | 806 | 1 FGR3_HUMAN | P22607 homo sapien |
| 36 | 326.5 | 25.1 | 557 | 1 SRC_AVIS1  | P14085 avian sarco |
| 37 | 325.5 | 25.0 | 533 | 1 FYN_CHICK  | Q05876 gallus gall |
| 38 | 324.5 | 25.0 | 535 | 1 SRC_HUMAN  | P12931 homo sapien |
| 39 | 324.5 | 25.0 | 568 | 1 SRC_MOUSE  | P14084 avian sarco |
| 40 | 324.5 | 25.0 | 801 | 1 FGR3_MOUSE | Q61851 mus musculu |
| 41 | 324   | 24.9 | 535 | 1 YRK_CHICK  | Q02977 gallus gall |
| 42 | 323   | 24.8 | 812 | 1 FGR1_XENLA | P22182 xenopus lae |
| 43 | 322.5 | 24.8 | 450 | 1 CSK_RAT    | P32577 rattus norv |
| 44 | 321   | 24.7 | 528 | 1 YES_AVISY  | P00527 avian sarco |
| 45 | 321   | 24.7 | 541 | 1 YES_CHICK  | P09324 gallus gall |

ALIGNMENTS

RESULT 1  
M3K9\_HUMAN  
ID M3K9\_HUMAN STANDARD; PRT; 394 AA.  
AC P80192;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Mitogen-activated protein kinase kinase kinase 9 (EC 2.7.1.-) (Mixed  
DE lineage kinase 1) (Fragment).  
GN MAP3K9 OR MLK1 OR PRK1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon epithelium;  
RX MEDLINE=93238756; PubMed=8477742;  
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;  
RT "Identification of a new family of human epithelial protein kinases  
RT containing two leucine/isoleucine-zipper domains.";  
RL Eur. J. Biochem. 213:701-710(1993).  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF  
CC COLONIC, BREAST AND OESOPHAGEAL ORIGIN.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC MAP KINASE KINASE KINASE SUBFAMILY.  
CC PIR; S32467; S32467.  
DR PIR; JU0229; JU0229.  
DR HSSP; P12931; 1FMK.  
DR GENE; HGNC:6861; MAP3K9.  
DR MIT; 600136;  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00219; TYK1; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;  
KW ATP-binding.  
FT NON TER 1 1  
FT DOMAIN 3 271 PROTEIN KINASE.  
FT NP BIND 9 17 ATP (BY SIMILARITY).  
FT BINDING 30 30 ATP (BY SIMILARITY).  
FT ACT SITE 127 127 BY SIMILARITY.  
FT DOMAIN 289 310 LEUCINE-ZIPPER 1 (BY SIMILARITY).  
FT DOMAIN 324 345 LEUCINE-ZIPPER 2 (BY SIMILARITY).  
FT DOMAIN 354 368 ARG/LYS-RICH (BASIC).  
SQ SEQUENCE 394 AA; 44975 MW; DBE40B7D31047FD8 CRC64;  
Query Match 39.3%; Score 510.5; DB 1; Length 394;  
Best local Similarity 42.7%; Pred. No. 5e-34;  
Matches 108; Conservative 44; Mismatches 76; Indels 25; Gaps 4;

QY 1 YRAKWIISQDEAVK-----KLKIEKEAEILSVLSHRNIIQFYGVILEPPNY 48  
DB 18 YRAFWIGD--EVAVKARHDPDEDISQTIENVRQEAFLFAMLKHPNIIALRGVCLKEPNL 75  
QY 49 GIVTEYASLSGLDYINSNRSEMDMDHIMTWATDVAKGMHYLHMEAPVKVIRDLKSRN 108  
DB 76 CLVMEFARGGGLNRLVSGKR---IPPDILVNMAVOIARGMNYLHDEAIVPIIHRDLKSSN 132  
QY 109 VVI-----AADGVLKICDFGASRFHNHTTHMSLVGTFPMMAPEVIOQLPVSETCDTY 160  
DB 133 ILILQKVENGLDSNKLKLTDFGLAREWHRTTKMSAAGTYAMMAPEVIRASMFSGSDVW 192  
QY 161 SYGVVLWEMLTREVPFKGLEGLQVAVLVEKNERLTTPSCPRSFAPLHQCWEADAKR 220  
DB 193 SYGVLLWELLTGEVFPFRGIDGRVAVGVAMNKLALPIPTCPPEPFAKLMEDCWNPDPSR 252  
QY 221 PSFKQIISILES 233  
DB 253 PSFTNILDQLTTI 265  
M3KA\_HUMAN STANDARD; PRT; 954 AA.  
ID M3KA\_HUMAN  
AC Q02779; Q12761; Q14871;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Mitogen-activated protein kinase kinase kinase 10 (EC 2.7.1.37)  
DE (Mixed lineage kinase 2) (Protein kinase MST).  
GN MAP3K10 OR MLK2 OR MST.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=96128179; PubMed=8536694;  
RA Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,  
RA Sutherland G.R., Simpson R.J.;  
RT "Complete nucleotide sequence, expression, and chromosomal  
RT localisation of human mixed-lineage kinase 2.";  
RL Eur. J. Biochem. 234:492-500(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=95249256; PubMed=7731697;  
RA Katoh M., Hirai M., Sugimura T., Terada M.;  
RT "Cloning and characterization of MST, a novel (putative)  
RT serine/threonine kinase with SH3 domain.";  
RL Oncogene 10:1447-1451(1995).  
RN [3]  
RP SEQUENCE OF 244-480 FROM N.A.  
RC TISSUE=Colon epithelium;  
RX MEDLINE=93238756; PubMed=8477742;  
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;  
RT "Identification of a new family of human epithelial protein kinases  
RT containing two leucine/isoleucine-zipper domains.";  
RL Eur. J. Biochem. 213:701-710(1993).  
RN [4]  
RP SEQUENCE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC -1- SIMILARITY: KINASE KINASE SUBFAMILY.  
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

CC -----  
DR EMBL; X90846; CAA62351.1; -.  
DR EMBL; Z48615; CAA88531.1; -.  
DR PIR; S32468; S32468.  
DR HSSP; P11362; 1FGK.  
DR Genew; HGNC:6849; MAP3K10.  
DR MIM; 600137; -.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00018; SH3; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00221; STYK; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
DR PROSITE; PS50011; PROTEIN KINASE\_DOM; 1.  
DR PROSITE; PS50002; SH3; 1.  
DR Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;  
KW ATP-binding; SH3 domain.  
FT DOMAIN 2 5 POLY-GLU.  
FT DOMAIN 16 81 SH3.  
FT NP\_BIND 98 360 PROTEIN KINASE.  
FT BINDING 104 112 ATP (BY SIMILARITY).  
FT ACT SITE 125 125 ATP (BY SIMILARITY).  
FT ACT SITE 222 222 BY SIMILARITY.  
FT DOMAIN 384 405 LEUCINE-ZIPPER 1 (BY SIMILARITY).  
FT DOMAIN 419 440 LEUCINE-ZIPPER 2 (BY SIMILARITY).  
FT DOMAIN 449 463 ARG/LYS-RICH (BASIC).  
FT CONFLICT 462 464 SRL -> AV (IN REF. 2).  
FT CONFLICT 465 480 LKREGGSHISLPSGF -> AQAGRRQHPALWL (IN  
FT REF. 3).  
FT CONFLICT 471 471 G -> S (IN REF. 2).  
FT CONFLICT 807 807 G -> R (IN REF. 2).  
FT CONFLICT 818 818 V -> A (IN REF. 2).  
SQ SEQUENCE 954 AA; 103623 MW; 538F4AAA559B0ABA CRC64;  
Query Match 37.7%; Score 490.5; DB 1; Length 954;  
Best Local Similarity 42.3%; Pred. No. 5.7e-32;  
Matches 113; Conservative 35; Mismatches 92; Indels 27; Gaps 6;  
QY 1 YRAKWIISQDEAVKKL-LKIEK-----EAILSVLSHRNIIQFYGVILEPPNY 48  
DB 113 YRALW--RGEEVAVKARLDPEKDPVTAEQVCEARLFGALQHPNIIALRGACLNPPHL 170  
QY 49 GIVTEYASLSGLDYINSNRSEMDMDHIM-TWATDVAKGMHYLHMEAPVKVIRDLKSR 107  
DB 171 CLVMEYARGALSRVLAGRRV-----PHVLVNMAVOIARGMNYLHNDAPVPIIHRDLKSI 226  
QY 108 NVVI-----AADGVLKICDFGASRFHNHTTHMSLVGTFPMMAPEVIOQLPVSETCDT 159  
DB 227 NILLEAIENHNLAIVLKITDFGLAREWHRTTKMSAAGTYAMMAPEVIRLSFSKSDV 286  
QY 160 SYGVVLWEMLTREVPFKGLEGLQVAVLVEKNERLTTPSCPRSFAPLHQCWEADAKK 219  
DB 287 WSFVLLWELLTGEVFPFRGIDGRVAVGVAMNKLTLPIPTCPPEPFAKLMEDCWDPDPHG 346  
QY 220 PSFKQIISILESMSNDTSLPDKNSF 246  
DB 347 RPDFGSILKRLVIEQSALEQMPLESF 373  
RESULT 3  
M3KC\_HUMAN STANDARD; PRT; 859 AA.  
ID M3KC\_HUMAN  
AC Q12852;  
DT 16-OCT-2001 (Rel. 40, Created)

```

DT 16-OCT-2001 (Rel. 40, last sequence update)
DE 16-OCT-2001 (Rel. 40, last annotation update)
DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)
DE (Leucine-zipper protein kinase) (ZPK).
GN MAP3K12 OR ZPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Teratocarcinoma;
RX MEDLINE=94311945; PubMed=8037767;
RA Reddy U.R., Pleasure D.;
RT "Cloning of a novel putative protein kinase having a leucine zipper
RL domain from human brain.";
CC Biochem. Biophys. Res. Commun. 202:613-620(1994).
CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
CC similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in brain and kidney.
CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, U07358; AAA67343.1; -.
DR HSSP; P12931; 1FMK.
DR Genew; HGNC:6851; MAP3K12.
DR MIM; 600447; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Magnesium; Membrane.
FT DOMAIN 125 366 PROTEIN KINASE.
FT NP_BIND 131 139 ATP (BY SIMILARITY).
FT BINDING 152 152 ATP (BY SIMILARITY).
FT ACT_SITE 236 236 BY SIMILARITY.
FT DOMAIN 665 668 POLY-PRO.
FT DOMAIN 720 725 POLY-GLU.
SQ SEQUENCE 859 AA; 93188 MW; 0E5209792C5C6F05 CRC64;
Query Match 37.7%; Score 490; DB 1; Length 859;
Best Local Similarity 44.3%; Pred. No. 5.6e-32;
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;
QY 10 KEVAVKLLKIEKEAEI--LSVLSHRNIIQFYGVILEPPNYGIIVTEYASLGSLYDINSN 67
DB 147 EEVAVKVRDL-KETDIKLRKLRKHPNIIITFKGVCCTQAPCYCILMEFCAGDYLEVLRAQ 205
QY 68 RSEEMDMHIMTWATVAKGMHYLHMEAPVKVIRHDLKSRNVIAADGVLKICDFGASR- 126
DB 206 RPTVPSL-LVDWSMGIAIGSMNYLHLH--KIIRDLKSPNMLITYDDVVKISDFTSKE 260

```

```

QY 127 FHNHTHMSLVGTFFPMNAPEVIQSLPVSETCDITYSGVVLWEMLTREVPFKGLEQVAM 186
DB 261 LSDKSTKMSFAGTVANNAPEVIRNEPVSEKVDIWSFGVLLWELLTGEIPYKDVDSATIW 320
QY 187 LVEKNERLTIPSSCPSPFAELLHQCEWADAKRPSFKQIISLESMSNDT-SLPDK 242
DB 321 GVGNSLSLHPVPSSCPDGFKILLRQCWNKSKPRNPSFRQILLHLDIASADVLTSPQE 377
RESULT 4
M3KC_MOUSE STANDARD; PRT; 888 AA.
AC Q60700; P70286;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)
DE (Leucine-zipper protein kinase) (ZPK) (Dual leucine zipper bearing
DE kinase) (DLK).
GN MAP3K12 OR ZPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Brain;
RX MEDLINE=95074107; PubMed=7983011;
RA Holzman L.B., Merritt S.E., Fan G.;
RT "Identification, molecular cloning, and characterization of dual
RT leucine zipper bearing kinase. A novel serine/threonine protein kinase
RT that defines a second subfamily of mixed lineage kinases.";
RL J. Biol. Chem. 269:30808-30817(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR X Swiss Webster; TISSUE=Brain;
RX MEDLINE=96365388; PubMed=8769565;
RA Blouin R., Beaudoin J., Bergeron P., Nadeau A., Grondin G.;
RT "Cell-specific expression of the ZPK gene in adult mouse tissues.";
RL DNA Cell Biol. 15:631-642(1996).
RN [3]
RP PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.
RX MEDLINE=96279269; PubMed=8663324;
RA Mata M., Merritt S.E., Fan G., Holzman L.B.;
RT "Characterization of dual leucine zipper-bearing kinase, a mixed
RT lineage kinase present in synaptic terminals whose phosphorylation
RT state is regulated by membrane depolarization via calcineurin.";
RL J. Biol. Chem. 271:16888-16896(1996).
CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
CC -1- TISSUE SPECIFICITY: Expressed in brain, kidney, lung, heart,
CC testis, gastrointestinal tract, stomach, liver and pancreas.
CC Within the nervous system, predominantly expressed in neurons and
CC enriched in synaptic terminals.
CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```



```

DR EMBL; U14636; AAA57280.1; -.
DR EMBL; U23789; AAB17123.1; -.
DR HSSP; P12931; 1FMK.
DR MGSI; 1346881; Map3k12.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR KX Transferase; Serine/threonine-protein kinase; ATP-binding;
KW phosphorylation; Magnesium; Membrane.
FT DOMAIN 158 399 PROTEIN_KINASE.
FT NP_BIND 164 172 ATP (BY SIMILARITY).
FT BINDING 185 185 ATP.
FT ACT_SITE 269 269 BY SIMILARITY.
FT DOMAIN 56 62 POLY-GLY.
FT DOMAIN 668 671 POLY-PRO.
FT DOMAIN 698 701 POLY-PRO.
FT DOMAIN 753 758 POLY-GLU.
FT MUTAGEN 185 185 K->A: NO CATALYTIC ACTIVITY.
FT MUTAGEN 192 192 E->A: NO CHANGE.
FT CONFLICT 18 18 LV->A (IN REF. 2).
FT CONFLICT 28 29 KL->N (IN REF. 2).
FT CONFLICT 382 382 S->T (IN REF. 2).
FT CONFLICT 494 495 EQ->DE (IN REF. 2).
FT CONFLICT 517 517 N->D (IN REF. 2).
FT CONFLICT 794 794 E->G (IN REF. 2).
SQ SEQUENCE 888 AA; 96083 MW; CFECCF1D34F889ABB CRC64;

```

|                       |                  |                    |            |             |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match           | 37.7%            | Score 490;         | DB 1;      | Length 888; |
| Best Local Similarity | 44.3%            | Pred. No. 5.8e-32; |            |             |
| Matches 105;          | Conservative 43; | Mismatches 79;     | Indels 10; | Gaps 6;     |

[illegible]

|             |  |      |         |
|-------------|--|------|---------|
| RESULT 5    |  |      |         |
| M3KC_RAT    |  |      |         |
| ID_M3KC_RAT | STANDARD;  | PRT; | 888 AA. |
| AC          | Q63796;  |      |         |
| DT          | 16-OCT-2001 (Rel. 40, Created)                                       |      |         |
| DT          | 16-OCT-2001 (Rel. 40, Last sequence update)                          |      |         |
| DT          | 16-OCT-2001 (Rel. 40, Last annotation update)                        |      |         |
| DE          | Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)             |      |         |
| DE          | (MAPK-upstream kinase) (MUK).  |      |         |
| GN          | MAP3K12 OR MUK.  |      |         |
| OS          | Rattus norvegicus (Rat).   |      |         |
| OC          | Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;    |      |         |
| OC          | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus |      |         |
| OX          | NCBI_TaxID=10116;  |      |         |
| RN          | [1]  |      |         |
| RP          | SEQUENCE FROM N.A.   |      |         |
| RX          | MEDLINE=96226099; PubMed=8637721;                                    |      |         |
| RT          | Hirai S., Izawa M., Osada S., Spyrou G., Ohno S.,                    |      |         |
|             | "Activation of the JNK pathway by distantly related protein kinases, |      |         |

```

RT MEKK and MUK."
RL Oncogene 12:641-650(1996).
CC -I- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -I- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -I- COFACTOR: Magnesium.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
CC similarity).
CC -I- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated (By similarity).
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

|    |  |
|----|--|
| DR | EMBL; D49785; BAA08621.1; -.                               |
| DR | HSSP; P12931; 1FMK.  |
| DR | InterPro; IPR000719; Euk_pkinase.                          |
| DR | InterPro; IPR004040; STY_pkinase.                          |
| DR | InterPro; IPR002290; Ser_thr_pkinase.                      |
| DR | Pfam; PF00069; pkinase; 1.                                 |
| DR | ProDom; PD000001; Euk_pkinase; 1.                          |
| DR | SMART; SM00221; STYKc; 1.                                  |
| DR | PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.           |
| DR | PROSITE; PS00108; PROTEIN_KINASE_ST; 1.                    |
| DR | PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.                   |
| KW | Transferase; Serine/threonine-protein kinase; ATP-binding; |
| KW | Phosphorylation; Magnesium; Membrane.                      |
| FT | DOMAIN 158 399 PROTEIN KINASE.                             |
| FT | NP_BIND 164 172 ATP (BY SIMILARITY).                       |
| FT | BINDING 185 185 ATP (BY SIMILARITY).                       |
| FT | ACT_SITE 269 269 BY SIMILARITY.                            |
| FT | DOMAIN 56 62 POLY-GLY.                                     |
| FT | DOMAIN 668 671 POLY-PRO.                                   |
| FT | DOMAIN 698 701 POLY-PRO.                                   |
| FT | DOMAIN 753 758 POLY-GLU.                                   |
| SQ | SEQUENCE 888 AA; 96307 MW; 52AD964006BAE149 CRC64;         |

|                           |       |                    |            |             |
|---------------------------|-------|--------------------|------------|-------------|
| Query Match               | 37.5% | Score 488;         | DB 1;      | Length 888; |
| Best Local Similarity     | 44.3% | Pred. No. 8.4e-32; |            |             |
| Matches 105; Conservative | 42;   | Mismatches 80;     | Indels 10; | Gaps 6;     |

[illegible]

| RESULT 6   |             |                                 |
|------------|-------------|---------------------------------|
| CTRL_ARATH |             |                                 |
| ID         | CTRL_ARATH  | STANDARD; PRT; 821 AA           |
| AC         | Q05609;     |                                 |
| DT         | 01-NOV-1995 | (Rel. 32, Created)              |
| DT         | 01-NOV-1995 | (Rel. 32, Last sequence update) |

DT 16-OCT-2001 (Rel. 40, last annotation update)  
 DE Serine/threonine-protein kinase CTR1 (EC 2.7.1.37).  
 GN CTR1 OR AT5G03730 OR F17C15\_150.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosidia II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia; TISSUE=Seedling;  
 RX MEDLINE=93161417; PubMed=8431946;  
 RA Kieber J.J., Rothenberg M., Roman G., Feldmann K.A., Ecker J.R.;  
 RT "CTR1, a negative regulator of the ethylene response pathway in  
 RL Arabidopsis, encodes a member of the rat family of protein kinases.";  
 RN Cell 72:427-441(1993).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016721; PubMed=11130714;  
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,  
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,  
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,  
 RA Nakazaki N., Naito K., Okumura S., Shinpo S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,  
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,  
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,  
 RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,  
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,  
 RA Du H., Edwards J., Fryman J., Hakenberg B., Lamar E., Latreille P.,  
 RA Leonard S., Meyer R., Mulvaney E., Ozerisky P., Riley A., Strommatt C.,  
 RA Wagner-McPherson C., Wollam A., Yoakum M., Bell M., Dedhia N.,  
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,  
 RA Kirchoff K., Toch K., King L., Bahret A., Miller B., Marra M.,  
 RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,  
 RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,  
 RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,  
 RA Langham S.-A., McCullagh B., Robben J., Gzymonprez B., Zimmermann W.,  
 RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,  
 RA van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R.,  
 RA Weltzenegger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,  
 RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,  
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,  
 RA Schueller C., Zaccaria F., Mewes H.-W., Bevan M., Franz P.F.;  
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis  
 RL thaliana.";  
 Nature 408:823-826(2000).  
 -1- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE  
 PATHWAY.  
 -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 -1- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.  
 -1- MISCELLANEOUS: CTR1 MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,  
 RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED  
 COTYLEDON GROWTH IS IMPAIRED.  
 -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 MI/RAF SUBFAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; L08789; AAA32779.1; -;  
 DR EMBL; L08790; AAA32780.1; -;  
 DR EMBL; AL162506; CAB82938.1; -;  
 DR InterPro; IPR000719; Euk\_kinase.  
 DR InterPro; IPR004040; STY\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR Pfam; PF00069; kinase; 1.  
 DR ProDom; PD000001; Euk\_kinase; 1.

DR SMART; SMO0221; STYKC; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.  
 FT DOMAIN 65 69 POLY-GLY.  
 FT DOMAIN 135 141 POLY-GLY.  
 FT DOMAIN 551 809 PROTEIN KINASE.  
 FT NP BIND 557 565 ATP (BY SIMILARITY).  
 FT BINDING 578 578 ATP (BY SIMILARITY).  
 FT ACT SITE 676 676 BY SIMILARITY.  
 FT MUTAGEN 596 596 E->K: IN CTR1-4; EXHIBITS ETHYLENE-  
 FT TREATED PHENOTYPE.  
 FT MUTAGEN 694 694 D->E: IN CTR1-1; EXHIBITS ETHYLENE-  
 FT TREATED PHENOTYPE.  
 SQ SEQUENCE 821 AA; 90306 MW; 2922D3DCD0CC15BC CRC64;  
 Query Match 34.0%; Score 442; DB 1; Length 821;  
 Best Local Similarity 40.2%; Pred. No. 3.9e-28;  
 Matches 99; Conservative 41; Mismatches 90; Indels 16; Gaps 5;  
 QY 1 YRAKWSIQDKEVAVKLLKIE-----KEAELSVLSHRNIQFYGVILEPPNYGI 50  
 Db 566 HRAEWHGSD--VAVKIMEQDFHAERVNEFLREVAIMKRLRHPNIVLFMGAVTQPNLSI 623  
 QY 51 VTEYASLGSLYDYIN--SNRSEMDMDHMTWATDVAKGMHYLHMEAPVYTHRDLKSRNV 109  
 Db 624 VTEYLSRGSGLYRLHKSGAREQLDERRRLSMAYDVAKGMNYLHNRNP--PIVHRDLKSPNL 682  
 QY 110 VIAADGVLCIDFGASRFHNT--THMSLVGTFPWMAPEVIOSLPVSETCDTYSYGVLM 167  
 Db 683 LVDKKYTVKVCDFGLSRLKASTFLSSKSAAGTPEWMAPEVLRLDEPSNDSYSGVILW 742  
 QY 168 EMLTREVPPFKGLEGLQVAWLVEKNERLTIPSSCPRSFAELLHQCEADAKKRPSFKQII 227  
 Db 743 ELATLQOPWGNLNDPAQVAAVGFCKRLRLEPRNLNPQVAIIIEGCTNTEPWRKPSFATIM 802  
 QY 228 SILESM 233  
 Db 803 DLLRPL 808  
 RESULT 7  
 M3K7\_MOUSE STANDARD; PRT; 579 AA.  
 ID M3K7\_MOUSE  
 AC Q62073;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, last sequence update)  
 DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)  
 DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-  
 DE activated kinase 1).  
 GN MAP3K7 OR TAK1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96123277; PubMed=8533096;  
 RA Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,  
 RA Taniguchi T., Nishida E., Matsumoto K.;  
 RT "Identification of a member of the MAPKK family as a potential  
 RT mediator of TGF-beta signal transduction.";  
 RT Science 270:2008-2011(1995).  
 RL Science 270:2008-2011(1995).  
 -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.  
 CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B  
 CC ACTIVATION.  
 -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE KINASE SUBFAMILY.  
 -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D76446; BAA1184.1; -  
DR HSSP; P08631; IAD5.  
DR MGD; MGI:1346877; Map3K7.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00221; STYKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR Transferase; Serine/threonine-protein kinase; ATP-binding.  
KW DOMAIN 8 POLY-SER.  
FT NP BIND 36 291 PROTEIN\_KINASE.  
FT BINDING 42 50 ATP (BY SIMILARITY).  
FT ACT SITE 63 63 ATP (BY SIMILARITY).  
FT ACT SITE 156 156 BY SIMILARITY.  
SQ SEQUENCE 579 AA; 64227 MW; 97C8F6F3C8E283EE CRC64;

Query Match 30.4%; Score 395; DB 1; Length 579;  
Best Local Similarity 37.1%; Pred. No. 1.7e-24;  
Matches 91; Conservative 46; Mismatches 80; Indels 28; Gaps 11;

QY 2 RAKWISQDKEVAVKKLIEKEAE-----ILSVLSHRNIIQFYGVILEPPNYGIYT 52  
DB 52 KAKW--RAKDVAIK--QIESESEKAFIVELRLQSRVNHPIVVKLYGACINP--VCLVM 104  
QY 53 EYASLGSLVDYINSNRSEEM--DMDHIMTWATDVAKGMHYLMHAPVKVIHRDLKSRNV 109  
DB 105 EYAEAGSLYNVLHG--AEPLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNL 162  
QY 110 VIAADG-VLKICDFG-ASRFNHTTHMSLVGTFPMWAPVIOSLPVSETCDTYSYGVLM 167  
DB 163 LTVAGGTVLKICDFGTACDIQTHMTNNK--GSAAMWAPVEFEGSNYSEKCDVFSWGIIWM 220  
QY 168 EMLTREVPFKGLEG--LQVAVLVVEKNERLTIPSSCPRSFAELLHQCWEADAKRPSFKQ 225  
DB 221 EVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPSMEE 279  
QY 226 IISIL 230  
DB 280 IVKIM 284

RESULT 8

MAK7 HUMAN STANDARD; PRT; 606 AA.  
ID 043318; 043317; 043319;  
AC 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)  
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-  
DE activated kinase 1).  
GN MAP3K7 OR TAK1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.  
RC TISSUE=Lung;  
RX MEDLINE=98153801; PubMed=9480845;  
RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.;  
RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an

RT NF-kappa B-inducing kinase-independent mechanism.";  
RL Biochem. Biophys. Res. Commun. 243:545-549(1998).  
CC -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKs.  
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B  
CC ACTIVATION.  
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1A, 1B (SHOWN HERE) AND 1C; ARE  
CC PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC MAP KINASE KINASE SUBFAMILY.  
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION, STRONG, TO C21ORF7.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AB009357; BAA25026.1; -  
DR EMBL; AB009356; BAA25025.1; -  
DR EMBL; AB009358; BAA25027.2; -  
DR HSSP; P08631; IAD5.  
DR Genew; HGNC:6859; MAP3K7.  
DR MIM; 602614; -  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00221; STYKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
KW Alternative splicing.  
FT DOMAIN 8 14 POLY-SER.  
FT NP BIND 36 291 PROTEIN\_KINASE.  
FT BINDING 42 50 ATP (BY SIMILARITY).  
FT ACT SITE 63 63 ATP (BY SIMILARITY).  
FT ACT SITE 156 156 BY SIMILARITY.  
FT VARSPLIC 404 430 MISSING (IN ISOFORM 1A).  
FT VARSPLIC 509 518 PLAPCNSKE -> ARTSCRTGPG (IN ISOFORM 1C).  
FT VARSPLIC 519 606 MISSING (IN ISOFORM 1C).  
SQ SEQUENCE 606 AA; 67196 MW; 3D8F8147CD174013 CRC64;

Query Match 30.4%; Score 395; DB 1; Length 606;  
Best Local Similarity 37.1%; Pred. No. 1.7e-24;  
Matches 91; Conservative 46; Mismatches 80; Indels 28; Gaps 11;

QY 2 RAKWISQDKEVAVKKLIEKEAE-----ILSVLSHRNIIQFYGVILEPPNYGIYT 52  
DB 52 KAKW--RAKDVAIK--QIESESEKAFIVELRLQSRVNHPIVVKLYGACINP--VCLVM 104  
QY 53 EYASLGSLVDYINSNRSEEM--DMDHIMTWATDVAKGMHYLMHAPVKVIHRDLKSRNV 109  
DB 105 EYAEAGSLYNVLHG--AEPLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNL 162  
QY 110 VIAADG-VLKICDFG-ASRFNHTTHMSLVGTFPMWAPVIOSLPVSETCDTYSYGVLM 167  
DB 163 LTVAGGTVLKICDFGTACDIQTHMTNNK--GSAAMWAPVEFEGSNYSEKCDVFSWGIIWM 220  
QY 168 EMLTREVPFKGLEG--LQVAVLVVEKNERLTIPSSCPRSFAELLHQCWEADAKRPSFKQ 225  
DB 221 EVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPSMEE 279  
QY 226 IISIL 230  
DB 280 IVKIM 284

RESULT 9

ABL1\_HUMAN



ID ABL1 HUMAN STANDARD; PRT; 1130 AA.  
AC P00519; Q16133, Q13869; Q13870;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Proto-oncogene tyrosine-protein kinase ABL1 (EC 2.7.1.112) (p150)  
DE (C-ABL).  
GN ABL1 OR ABL OR JTK7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fibroblast;  
RX MEDLINE=90082420; PubMed=2687768;  
RA Fainstein E., Einat M., Gokkel E., Marcelle C., Croce C.M.,  
RA Gale R.P., Canaan E.;  
RT "Nucleotide sequence analysis of human abl and bcr-abl cDNAs.";  
RT Oncogene 4:1477-1481(1989).  
RN [2]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RC MEDLINE=87028219; PubMed=3021337;  
RX Shtivelman E., Lifshitz B., Gale R.P., Roe B.A., Canaan E.;  
RT "Alternative splicing of RNAs transcribed from the human abl gene and  
RT from the bcr-abl fused gene.";  
RT Cell 47:277-284(1986).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORMS IA AND IB).  
RC TISSUE=Lung carcinoma;  
RX MEDLINE=95394474; PubMed=7665185;  
RA Chisoe S.W., Bodenteich A., Wang Y.-F., Wang Y.-P., Burian D.,  
RA Clifton S.W., Crabtree J., Freeman A., Iyer K., Jian L., Ma Y.,  
RA McLaury H.-J., Pan H.-Q., Sathan O.H., Toth S., Wang Z., Zhang G.,  
RA Heisterkamp N., Groffen J., Roe B.A.;  
RT "Sequence and analysis of the human ABL gene, the BCR gene, and  
RT regions involved in the Philadelphia chromosome translocation.";  
RT Genomics 27:67-82(1995).  
RN [4]  
RP SEQUENCE OF 360-426 FROM N.A.  
RC MEDLINE=83245023; PubMed=6191223;  
RX Groffen J., Heisterkamp N., Reynolds F.H. Jr., Stephenson J.R.;  
RT "Homology between phosphotyrosine acceptor site of human c-abl and  
RT viral oncogene products.";  
RT Nature 304:167-169(1983).  
RN [5]  
RP SEQUENCE OF 27-40 FROM N.A.  
RC MEDLINE=88065859; PubMed=2825022;  
RX Fainstein E., Marcelle C., Rosner A., Canaan E., Gale R.P.,  
RX Drazek O., Smith S.D., Croce C.M.;  
RT "A new fused transcript in Philadelphia chromosome positive acute  
RT lymphocytic leukaemia.";  
RT Nature 330:386-388(1987).  
RN [6]  
RP SEQUENCE OF 825-845 FROM N.A.  
RC MEDLINE=94142331; PubMed=7545908;  
RX Inokuchi K., Futaki M., Dan K., Nomura T.;  
RT "Sequence analysis of the mutation at codon 834 and the sequence  
RT variation of codon 837 of c-abl gene.";  
RT Leukemia 8:343-344(1994).  
RN [7]  
RP STRUCTURE BY NMR OF SH2 DOMAIN.  
RC MEDLINE=92370689; PubMed=1505033;  
RX Overduin M., Rios C.B., Mayer B.J., Baltimore D., Cowburn D.;  
RT "Three-dimensional solution structure of the src homology 2 domain of  
RT c-abl.";  
RT Cell 70:697-704(1992).  
RN [8]  
RP STRUCTURE BY NMR OF SH2 DOMAIN.  
RC MEDLINE=93101588; PubMed=1281542;  
RX Overduin M., Mayer B.J., Rios C.B., Baltimore D., Cowburn D.;  
RT "Secondary structure of Src homology 2 domain of c-Abl by  
RT heteronuclear NMR spectroscopy in solution.";

RL Proc. Natl. Acad. Sci. U.S.A. 89:11673-11677(1992).  
RN [9]  
RP STRUCTURE BY NMR OF SH3 DOMAIN.  
RX MEDLINE=96131878; PubMed=8590002;  
RA Gosser Y.O., Zheng J., Overduin M., Mayer B.J., Cowburn D.;  
RT "The solution structure of Abl SH3, and its relationship to SH2 in  
RT the SH(32) construct.";  
RT Structure 3:1075-1086(1995).  
RN [10]  
RP 3D-STRUCTURE MODELING OF SH3 DOMAIN.  
RX MEDLINE=95199229; PubMed=7892170;  
RA Pisabarro M.T., Ortiz A.R., Serrano L., Wade R.C.;  
RT "Homology modeling of the Abl-SH3 domain.";  
RT Proteins 20:203-215(1994).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 57-218.  
RX MEDLINE=96398698; PubMed=8805596;  
RA Nam H.-J., Haseg W.G., Roberts T.M., Frederick C.A.;  
RT "Intramolecular interactions of the regulatory domains of the Bcr-Abl  
RT kinase reveal a novel control mechanism.";  
RT Structure 4:1105-1114(1996).  
RN [12]  
RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS) OF 64-121.  
RX MEDLINE=98365516; PubMed=9698566;  
RA Pisabarro M.T., Serrano L., Wilmanns M.;  
RT "Crystal structure of the abl-SH3 domain complexed with a designed  
RT high-affinity peptide ligand: implications for SH3-ligand  
RT interactions.";  
RT J. Mol. Biol. 281:513-521(1998).  
RN [13]  
RP CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
RC tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; IA (SHOWN HERE) AND IB, ARE  
CC PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
CC -1- DISEASE: PARTICIPATES IN A T(9;22)(Q34;Q11) CHROMOSOMAL  
CC TRANSLOCATION THAT PRODUCES A BCR-ABL ONCOGENE RESPONSIBLE FOR  
CC CHRONIC MYELOID LEUKEMIA (CML), ACUTE MYELOID LEUKEMIA (AML), AND  
CC ACUTE LYMPHOBLASTIC LEUKEMIA (ALL).  
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
CC SUBFAMILY.  
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
CC WWW="http://www.infobiogen.fr/services/chromocancer/Genes/ABL.html".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X16416; CAA34438.1; -;  
DR EMBL; M14752; AAA51561.1; -;  
DR EMBL; U07563; AAB60394.1; -;  
DR EMBL; U07563; AAB60393.1; -;  
DR EMBL; U07561; AAB60393.1; JOINED.  
DR EMBL; S69223; AAD14034.1; -;  
DR PIR; A25582; TVHUA.  
DR PDB; 1AB2; 31-JAN-94.  
DR PDB; 1ABL; 01-NOV-94.  
DR PDB; 2ABL; 04-SEP-97.  
DR PDB; 1AWO; 28-JAN-98.  
DR PDB; 1BBZ; 25-NOV-98.  
DR Genew; HGNC:76; ABL1.  
DR MIM; 189980; -;  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00017; SH2; 1.



```
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
KW Transferase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;
KW Phosphorylation; SH2 domain; SH3 domain; Chromosomal translocation;
KW 3D-structure; Alternative splicing.
FT DOMAIN 61 121 SH3.
FT DOMAIN 127 217 SH2.
FT DOMAIN 242 493 PROTEIN KINASE.
FT DOMAIN 605 609 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 18 22 POLY-SER.
FT DOMAIN 605 609 POLY-LYS.
FT DOMAIN 782 1019 PRO-RICH.
FT DOMAIN 897 903 POLY-PRO.
FT SITE 26 27 BREAKPOINT FOR TRANSLOCATION TO FORM
FT NP_BIND 248 256 BCR-ABL ONCOGENE.
FT BINDING 271 271 ATP (BY SIMILARITY).
FT ACT_SITE 363 363 ATP (BY SIMILARITY).
FT MOD_RES 393 393 BY SIMILARITY.
FT VARSPIC 1 26 MLEICLKVGCKSKGKLGSSSSCYCLE -> MCGQFGKVLGD
OQRPSLPALHFIKAGKKESSRHGGPHCNVEYEH (IN
ISOFORM 1B).
FT CONFLICT 140 140 L -> P (IN REF. 2).
FT CONFLICT 159 159 G -> S (IN REF. 2).
FT CONFLICT 424 425 AF -> GK (IN REF. 4).
FT CONFLICT 445 445 L -> R (IN REF. 2).
FT CONFLICT 459 459 E -> K (IN REF. 2).
FT CONFLICT 520 520 S -> T (IN REF. 2).
FT CONFLICT 719 719 A -> V (IN REF. 2).
FT CONFLICT 837 837 E -> G (IN REF. 3 AND 6).
FT CONFLICT 837 837 E -> W (IN REF. 2).
FT CONFLICT 863 863 G -> R (IN REF. 2).
FT CONFLICT 894 894 R -> K (IN REF. 2).
FT CONFLICT 917 919 SPS -> RPG (IN REF. 2).
FT CONFLICT 952 952 G -> A (IN REF. 2).
FT CONFLICT 967 968 OS -> HP (IN REF. 2).
FT CONFLICT 983 983 S -> LS (IN REF. 2).
FT CONFLICT 1022 1022 MISSING (IN REF. 2).
FT CONFLICT 1045 1045 R -> G (IN REF. 2).
FT CONFLICT 1103 1103 T -> S (IN REF. 2).
FT TURN 120 120

Query Match 28.4%; Score 369.5; DB 1; Length 1130;
Best Local Similarity 32.4%; Pred. No. 3.9e-22;
Matches 82; Conservative 55; Mismatches 101; Indels 15; Gaps 7;
```

```
DB 434 TYGMSYPYCIDLSQV-YELLEKDYRMERPEGCEPEKYVELMRACQWNPSPDRPSFAEIHQA 492
QY 230 LESMSNDTSLPDK 242
DB 493 FETMFQESSISDE 505

RESULT 10
ABL_MLVAB STANDARD; PRT; 746 AA.
ID - ABL_MLVAB
AC P00521;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase transforming protein ABL (EC 2.7.1.112).
GN V-ABL.
OS Abelson murine leukemia virus.
OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_TaxID=11788;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83221648; PubMed=6304726;
RA Reddy E.P., Smith M.J., Srinivasan A.;
RT "Nucleotide sequence of Abelson murine leukemia virus genome:
RT structural similarity of its transforming gene product to other onc
RT gene products with tyrosine-specific kinase activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:3623-3627(1983).
RN [2]
RP REVISIONS TO 588-746.
RA Reddy E.P., Smith M.J., Srinivasan A.;
RL Proc. Natl. Acad. Sci. U.S.A. 80:7372-7372(1983).
RN [3]
RP SEQUENCE OF 233-327 FROM N.A.
RX MEDLINE=83245023; PubMed=6191223;
RA Groffen J., Heisterkamp N., Reynolds F.H. Jr., Stephenson J.R.;
RT "Homology between phosphotyrosine acceptor site of human c-abl and
RT viral oncogene products.";
RL Nature 304:167-169(1983).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL
CC POLYPEPTIDE.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V01541; -, NOT ANNOTATED_CDS.
CC EMBL; K00010; AAA46470.1; -.
DR EMBL; A00627; TVMVGW.
DR HSP; P00519; 2ABL.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
```

KW Polypeptide; Tyrosine-protein kinase; Transferase; Oncogene;  
 SH2 domain.  
 FT DOMAIN 13 103 SH2.  
 RT 128 379 PROTEIN KINASE.  
 SQ SEQUENCE 746 AA; 81872 MW; B9072FF5FE9257 CRC64;  
 Query Match 28.3%; Score 368.5; DB 1; Length 746;  
 Best Local Similarity 32.4%; Pred. No. 2.9e-22;  
 Matches 82; Conservative 54; Mismatches 102; Indels 15; Gaps 7;  
 QY 1 YRAKWSQDKEVAVKL---LKIE---KEAELSVLSHRNIQFYGVILEPPNYGIATE 53  
 DB 143 YEGVWKYSLTVAVKTLKEDTMEVEEFLKEAAVWKEIKHPNLVQLGCTREPPIIITE 202  
 QY 54 YASLGLDYINSNRSEEMDMHMTWATDVAKGMYLHMEAPVKVIRDLKSRNVIAA 113  
 DB 203 FMTYGNLDYLRNCRQEVSAVLLVMTQISSAMEYLEKK---NFIHDLARNCLVGE 259  
 QY 114 DGLKICDPGASRFNHTTMSLVGT-PP--WMAPEVIOQLPVSECTDYSYGVVLEWL 170  
 DB 260 NHLVAVADFGSLRLMTGDTYTAHAGAKFPIKWTAPESLAYNKFSLKSDVWAFGLWIEIA 319  
 QY 171 TRFV-PRKLEGLQVAVLVVEKNERLTIPSSCPSPFAELLHQEWADAKRPSFKQIISI 229  
 DB 320 TYGMSFYPGIDLSQV-YELLEKDYMERBECPEKVELMRACQWNPSPDRPSFAEIHQA 378  
 QY 230 LESMSNDTSLPDK 242  
 DB 379 FETMFQESSISDE 391

RESULT 11  
 ID ABL1\_MOUSE STANDARD; PRT; 1123 AA.  
 AC P00520; Q61252; Q61253; Q61254; Q61255; Q61256; Q61257; Q61258;  
 AC Q61259; Q61260; Q61261; P97896;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Proto-oncogene tyrosine-protein kinase ABL1 (BC 2.7.1.112) (p150)  
 DE (c-ABL).  
 GN ABL1 OR ABL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 SEQUENCE FROM N.A.  
 TISSUE=Testis;  
 MEDLINE=88068561; PubMed=3317402;  
 RA Oppi C., Shore S.K., Reddy E.P.;  
 RT "Nucleotide sequence of testis-derived c-abl cDNAs: implications for  
 RT testis-specific transcription and abl oncogene activation.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 84:8200-8204(1987).  
 RN [2]  
 SEQUENCE OF 1-187 FROM N.A. (ISOFORMS I TO IV).  
 RX MEDLINE=95394474; PubMed=7665185;  
 RA Chisoso S.L., Bodenteich A., Wang Y.-F., Wang Y.-P., Burian D.,  
 RA Clifton S.W., Crabtree J., Freeman A., Iyer K., Jian L., Ma Y.,  
 RA McLaury H.-J., Pan H.-Q., Sarhan O.H., Toch S., Wang Z., Zhang G.,  
 RA Helsterkamp N., Groffen J., Roe B.A.;  
 RT "Sequence and analysis of the human ABL gene, the BCR gene, and  
 RT regions involved in the Philadelphia chromosome translocation.";  
 RT Genomics 27:67-82(1995).  
 RN [3]  
 SEQUENCE OF 85-182 FROM N.A.  
 RX MEDLINE=84106840; PubMed=6319018;  
 RA Wang J.-Y., Ledley F., Goff S., Lee R., Groner Y., Baltimore D.;  
 RT "The mouse c-abl locus: molecular cloning and characterization.";  
 RL Cell 36:349-356(1984).  
 RN [4]  
 ALTERNATIVE SPLICING.  
 RX MEDLINE=88202920; PubMed=3283651;

RA Bernards A., Paskind M., Baltimore D.;  
 RT "Four murine c-abl mRNAs arise by usage of two transcriptional  
 RT promoters and alternative splicing.";  
 RL Oncogene 2:297-304(1988).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 61-121.  
 RX MEDLINE=95393198; PubMed=7664083;  
 RA Musacchio A., Saracino M., Wilmanns M.;  
 RT "High-resolution crystal structures of tyrosine kinase SH3 domains  
 RT complexed with proline-rich peptides.";  
 RL Nat. Struct. Biol. 1:546-551(1994).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; THE MYRISTOYLATED C-ABL PROTEIN  
 CC WAS REPORTED TO BE NUCLEAR.  
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; I (SHOWN HERE), II, III AND  
 CC IV; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -----  
 CC EMBL: J02995; AAA88241.1; -;  
 CC EMBL: U14721; AAB60451.1; -;  
 CC EMBL: U14720; AAB60451.1; JOINED.  
 CC EMBL: U14721; AAB60450.1; -;  
 CC EMBL: U14720; AAB60450.1; JOINED.  
 CC EMBL: U14721; AAB60448.1; -;  
 CC EMBL: U13835; AAB60448.1; JOINED.  
 CC EMBL: U13835; AAB60449.1; -;  
 CC EMBL: U13835; AAB60449.1; JOINED.  
 CC EMBL: X07539; CAA30411.1; -;  
 CC EMBL: X07539; CAA30412.1; -;  
 CC EMBL: X07540; CAA30413.1; -;  
 CC EMBL: X07541; CAA30414.1; -;  
 CC EMBL: M12263; AAA37136.1; -;  
 CC EMBL: M12264; AAA37137.1; -;  
 CC EMBL: M12265; AAA37138.1; -;  
 CC EMBL: M12266; AAA37134.1; -;  
 CC EMBL: K03228; AAA37135.1; -;  
 CC PIR; A00626; A00626.  
 CC PIR; A39962; A39962.  
 CC PDB; 1ABQ; 15-OCT-95.  
 CC PDB; 1ABQ; 15-OCT-95.  
 CC MGI; MGI:87859; Ab11.  
 CC InterPro; IPR000719; Euk\_pkinase.  
 CC InterPro; IPR000980; SH2.  
 CC InterPro; IPR001452; SH3.  
 CC InterPro; IPR001245; Tyr\_pkinase.  
 CC Pfam; PF00017; SH2; 1.  
 CC Pfam; PF00018; SH3; 1.  
 CC Pfam; PF00069; pkinase; 1.  
 CC PRINTS; PR00401; SH2DOMAIN.  
 CC PRINTS; PR00109; TYRKINASE.  
 CC ProDom; PD000001; Euk\_pkinase; 1.  
 CC ProDom; PD000066; SH3; 1.  
 CC ProDom; PD000093; SH2; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC PROSITE; PS50001; SH2; 1.  
 CC PROSITE; PS50002; SH3; 1.  
 KW Transferase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;  
 KW Phosphorylation; SH2 domain; SH3 domain; Alternative splicing;

KW Nuclear protein; 3D-structure.  
FT DOMAIN 61 121 SH3.  
FT DOMAIN 127 217 SH2.  
FT DOMAIN 242 493 PROTEIN KINASE.  
FT DOMAIN 605 609 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
FT DOMAIN 18 22 POLY-SER.  
FT DOMAIN 605 609 POLY-LYS.  
FT DOMAIN 804 1012 PRO-RICH.  
FT DOMAIN 891 897 POLY-PRO.  
FT NP BIND 248 256 ATP (BY SIMILARITY).  
FT BINDING 271 271 ATP (BY SIMILARITY).  
FT ACT SITE 363 363 BY SIMILARITY.  
FT MOD\_RES 393 393 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT VARSPPLIC 1 26 MLEICLVGCKSKKGLSSSSCYCLE -> MSQRWYTKCR  
FT VARSPPLIC 1 26 MLEICLVGCKSKKGLSSSSCYCLE -> MISFDLSDLE  
FT VARSPPLIC 1 26 HLKLVLDV (IN ISOFORM II).  
FT VARSPPLIC 1 26 MLEICLVGCKSKKGLSSSSCYCLE -> MGQPGKVLGD  
FT VARSPPLIC 1 26 QRRPSLPALHFIKKGAKGRDSSRHGPHCNVFEH (IN  
CONFLICT 184 187 LYVS -> VGDW (IN REF. 2).  
SEQUENCE 1123 AA; 122676 MW; 284F0830644AFD8F CRC64;  
Query Match 28.3%; Score 368.5; DB 1; Length 1123;  
Best Local Similarity 32.4%; Pred. No. 4.7e-22;  
Matches 82; Conservative 54; Mismatches 102; Indels 15; Gaps 7;  
QY 1 YRAKWSQDKEVAVKL---LKIE---KEAELISVLSHRNIIQFYGVILEPPNYGIIVT 53  
DB 257 YEGVWKYSLTVAVKTLKEDTMEVEEFLKEAAVMKEIKHPNLVQLLGVCTREPPFYIITE 316  
QY 54 YASLGSLDYDINSNRSEEMDMHMTWATDVAKGMHYLMEAPVKVIHRDLKSRNVIAA 113  
DB 317 FMTYGNLLDYLRECNROEVSAYVLLYMATQISSAMEYLEKK--NFIHRDLAARNCLVGE 373  
QY 114 DGVLKICDFGASRFHNHTHMSLVGT-FP--WMAPEVIQSLPVSETCDTYSYGVVLMEM 170  
DB 374 NHLVYKVADEGLSRLMTGDTYTAHAGAKPEIKWTAPESLAVNKFISKSDVMAFGVLLMEIA 433  
QY 171 TREV-PEKGLEGLQVAMLVVEKNERLITPSSCPRSFAELLHQCEADAKRPSFKQIISI 229  
DB 434 TYGMSYPGIDLSQV-YELLEKDYMERPEGCEKVELMRAQWNPSPDRPSFAELHQA 492  
QY 230 LESMSNDTSLPDK 242  
DB 493 FETMFQESSISDE 505

CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
CC SUBFAMILY.  
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
CC -1- DATABASE: NAME=Atlas Genet. Cytoenet. Oncol. Haematol.;  
WWW="http://www.infobiogen.fr/services/chronocancer/Genes/ABL2ID226.html".  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
CC EMBL; M35296; AAA35553.1; -.  
CC HSSP; P00519; 1BBZ.  
CC Genew; HGNC:77; ABL2.  
CC MIM; 164690; -.  
CC InterPro; IPR000719; Euk\_pkinase.  
CC InterPro; IPR000980; SH2.  
CC InterPro; IPR001452; SH3.  
CC InterPro; IPR001245; Tyr\_pkinase.  
CC Pfam; PF00017; SH2; 1.  
CC Pfam; PF00018; SH3; 1.  
CC Pfam; PF00069; kinase; 1.  
CC PRINTS; PR00401; SH2DOMAIN.  
CC PRINTS; PR00109; TYRKINASE.  
CC Prodom; PD000001; Euk\_pkinase; 1.  
CC Prodom; PD000066; SH3; 1.  
CC Prodom; PD000093; SH2; 1.  
CC SMART; SM00252; SH2; 1.  
CC SMART; SM00326; SH3; 1.  
CC SMART; SM00219; TYRK; 1.  
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
CC PROSITE; PS50001; SH2; 1.  
CC PROSITE; PS50002; SH3; 1.  
CC Transferrase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;  
KW Phosphorylation; SH2 domain; SH3 domain; Alternative splicing.  
FT DOMAIN 107 167 SH2.  
FT DOMAIN 173 263 PROTEIN KINASE.  
FT DOMAIN 288 539 POLY-SER.  
FT DOMAIN 561 564 POLY-SER.  
FT DOMAIN 658 660 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
FT DOMAIN 732 739 POLY-GLY.  
FT DOMAIN 843 1055 PRO-RICH.  
FT DOMAIN 984 988 POLY-PRO.  
FT NP BIND 294 302 ATP (BY SIMILARITY).  
FT BINDING 317 317 ATP (BY SIMILARITY).  
FT ACT SITE 409 409 BY SIMILARITY.  
FT MOD\_RES 439 439 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT VARSPPLIC 1 73 MGQVGRVGEAPGLQOPQPRGIGSSAARPSGRRDPAGRT  
TETGFNFTQHDHFAASCVDEFGDKTGSSP -> NVLGT  
VLPNTYGRDQDTSLCCLCTEASESALPDLT (IN  
ISOFORM IA).  
SQ SEQUENCE 1182 AA; 128343 MW; ED93869BC2B14FAA CRC64;  
Query Match 28.1%; Score 365.5; DB 1; Length 1182;  
Best Local Similarity 35.0%; Pred. No. 8.6e-22;  
Matches 89; Conservative 47; Mismatches 101; Indels 17; Gaps 9;  
QY 1 YRAKWSQDKEVAVKL---LKIE---KEAELISVLSHRNIIQFYGV-ILEPPNYGIIVT 52  
DB 303 YEGVWKYSLTVAVKTLKEDTMEVEEFLKEAAVMKEIKHPNLVQLLGVCTLEPPFY-IVT 361  
QY 53 YASLGSLDYDINSNRSEEMDMHMTWATDVAKGMHYLMEAPVKVIHRDLKSRNVIA 112  
DB 362 EYMPYGNLLDYLRECNREEVTAVVLVLYMATQISSAMEYLEKK--NFIHRDLAARNCLVG 418  
QY 113 ADGVLKICDFGASRFHNHTHMSLVGT-FP--WMAPEVIQSLPVSETCDTYSYGVVLMEM 169

```
Db 419 ENHVKVADFGISRLMTGDTYTAHAGAKFPKWTAPESLAYNTFSIKSDVWAFGLLWEI 478
Qy 170 LTREV-PFKGLEGLQVAMLVKERNLTIPTSSCPSPFAELLHQCWEADAKRPSFKQIIS 228
Db 479 ATYGMSPPYPGIDLSQV-YDLEKGYRMEQEGCPKPYELMRACWKWSPADRPSPFAETHQ 537
Qy 229 ILESMSNDTSLPDK 242
Db 538 AFETMFHDSISEE 551

RESULT 13
ABL_DROME STANDARD; PRT; 1520 AA.
ID ABL_DROME STANDARD; PRT; 1520 AA.
AC P00522;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DT Tyrosine-protein kinase Abl (EC 2.7.1.112) (D-ash).
OC ABL OR DASH OR ABL-1.
OC Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=88174728; PubMed=2832740;
RA Henkemeyer M.J., Bennett R.L., Gertler F.B., Hoffmann F.M.;
RT "DNA sequence, structure, and tyrosine kinase activity of the
RT Drosophila melanogaster Abelson proto-oncogene homolog.";
RL Mol. Cell. Biol. 8:843-853(1988).
RN [2]
RP SEQUENCE OF 374-648 FROM N.A.
RP MEDLINE=84082064; PubMed=6317185;
RA Hoffmann F.M., Fresco L.D., Hoffman-Falk H., Shilo B.-Z.;
RT "Nucleotide sequences of the Drosophila src and abl homologs:
RT conservation and variability in the src family oncogenes.";
RL Cell 35:393-401(1983).
RN [3]
RP FUNCTION.
RP MEDLINE=98298928; PubMed=9635189;
RA Loureiro J., Pelzer M.;
RT "Roles of Armadillo, a Drosophila catenin, during central nervous
RT system development.";
RL Curr. Biol. 8:622-632(1998).
CC -1- FUNCTION: ARM AND ABL PROTEINS FUNCTION COOPERATIVELY AT ADHERENS
CC JUNCTIONS IN BOTH THE CNS AND EPIDERMIS.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M19692; AAA28934.1; -
CC EMBL; M19690; AAA28934.1; JOINED.
CC EMBL; M19691; AAA28934.1; JOINED.
CC EMBL; K01042; AAA28443.1; -
CC PIR; A28128; TVPFA.
CC HSSP; P00519; 1AB2.
CC FlyBase; FBgn0000017; Abl.
CC InterPro; IPR000719; Euk_kinase.
```

```
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00069; kinase; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_kinase.
DR ProDom; PD000066; SH3; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; phosphorylation;
KW SH2 domain; SH3 domain.
FT DOMAIN 204 265 SH3.
FT DOMAIN 271 363 SH2.
FT DOMAIN 388 644 PROTEIN_KINASE.
FT NP_BIND 394 402 ATP (BY SIMILARITY).
FT BINDING 417 417 ATP (BY SIMILARITY).
FT ACT_SITE 509 509 BY SIMILARITY.
FT MOD_RES 539 539 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 374 377 LSPE -> ASAQ (IN REF. 2).
FT CONFLICT 645 648 ESS1 -> VGDV (IN REF. 2).
SQ SEQUENCE 1520 AA; 161836 MW; AD6A5060579FAD7B CRC64;
```

Query Match 28.0%; Score 364.5; DB 1; Length 1520;  
Best Local Similarity 34.1%; Pred. No. 1.4e-21;  
Matches 86; Conservative 47; Mismatches 104; Indels 15; Gaps 7;

```
Qy 1 YRAKWSIQDKEVAVKKL-----LK-IEKEAELISVLSHRNIIQFYGVILEPPNYGIYTE 53
Db 403 YEAVWKRYGNTVAAYTLKEDTMAKDFLEEAIMKEMKHPNLVQLIGVCTREBPFIYTE 462
Qy 54 YASLGSLYDYINSNRSEMDMDHIMTWATDVAKGMHYLMEAPVKVIHRDLKSRNVIAA 113
Db 463 FMSHGNLLDFLRSGRETLDAVALLYMATQIASGMSYLESR---NYIHRDLAARNCLVGD 519
Qy 114 DGLKICDFGASRFNHTTMSLVGT-FP--WNAPEVIQSLPVSETCDTYSYGVLMWML 170
Db 520 NKLVKVADFGIARLNRDDTYTAHAGAKFPKWTAPBGLAYNKFSTKSDVWAFGLLWEIA 579
Qy 171 TREV-PFKGLEGLQVAMLVKERNLTIPTSSCPSPFAELLHQCWEADAKRPSFKQIISI 229
Db 580 TYGMSPPYPAID-LTDVYHKLDKGYRMERPPGCPPEVYDLMRQCWQWDATDRPTFSIHHA 638
Qy 230 ILESMSNDTSLPD 241
Db 639 LEHMFQESSITE 650
```

```
RESULT 14
KYK1 DICDI STANDARD; PRT; 1584 AA.
ID KYK1 DICDI STANDARD; PRT; 1584 AA.
AC P18160;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) (Tyrosine-
DE protein kinase 1).
GN PYKA OR SPLA OR DPYK1.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=JH10;
```



RX MEDLINE=97053827; PubMed=8898241;  
RA Nickolls G.H., Oshervov N., Loomis W.F., Spudich J.A.;  
RT "The Dictyostelium dual-specificity kinase splA is essential for  
RT spore differentiation.";  
RL Development 122:3295-3305(1996).  
[2]  
RP SEQUENCE OF 1248-1584 FROM N.A.  
RX MEDLINE=90287147; PubMed=1972546;  
RA Tan J.L., Spudich J.A.;  
RT "Developmentally regulated protein-tyrosine kinase genes in  
RT Dictyostelium discoideum.";  
RL Mol. Cell. Biol. 10:3578-3583(1990).  
CC -1- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK  
CC DURING THE MOUND STAGE OF MORPHOGENESIS.  
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.  
CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U32174; AAB41125.1; -  
DR EMBL; M33785; AAA33202.1; -  
DR PIR; A35670; A35670.  
DR DictyDb; DD03010; pykA.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR001660; SAM.  
DR InterPro; IPR003878; SPRY\_domain.  
DR InterPro; IPR003877; SPRY\_receptor.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00536; SAM; 1.  
DR Pfam; PF00622; SPRY; 3.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SMO0454; SAM; 1.  
DR SMART; SMO0449; SPRY; 3.  
DR SMART; SMO0221; STYK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS50105; SAM\_DOMAIN; 1.  
RT Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
FT DOMAIN 908 972 SAM.  
FT DOMAIN 403 420 POLY-ASN.  
FT DOMAIN 428 435 POLY-THR.  
FT DOMAIN 449 480 POLY-ASN.  
FT DOMAIN 483 491 POLY-ASN.  
FT DOMAIN 494 508 POLY-ASN.  
FT DOMAIN 512 532 POLY-ASN.  
FT DOMAIN 596 600 POLY-ASN.  
FT DOMAIN 808 811 POLY-PHE.  
FT DOMAIN 1026 1029 POLY-SER.  
FT DOMAIN 1195 1210 POLY-ASN.  
FT DOMAIN 1215 1220 POLY-GLN.  
FT DOMAIN 1224 1233 POLY-GLN.  
FT DOMAIN 1266 1274 POLY-PRO.  
FT DOMAIN 1289 1561 PROTEIN\_KINASE.  
FT NP\_BIND 1295 1303 ATP (BY SIMILARITY).  
FT BINDING 1316 1316 ATP (BY SIMILARITY).  
FT ACT\_SITE 1417 1417 BY SIMILARITY.  
FT CONFLICT 1248 1248 D -> R (IN REF. 2).  
FT CONFLICT 1435 1435 V -> L (IN REF. 2).  
SQ SEQUENCE 1584 AA; 174304 MW; 5D1589458D8E01E3 CRC64;

Query Match 27.3%; Score 355.5; DB 1; Length 1584;

Best Local Similarity 34.1%; Pred. No. 7.8e-21;  
Matches 93; Conservative 47; Mismatches 96; Indels 37; Gaps 8;  
QY 2 RAKWISQDKEVAVK-----KLKIEKAEILSVLSHRNIIQFYVILE--PPNY 48  
DB 1305 RGYW--RETDVAIKIIRYDQFKTKSLVMPQNEVGILSKLRHPNVQFLGACTAGEDDH 1362  
QY 49 GIVTEYASLSLYDINSNRSEEMDMHI-MTWTADVAKGMHYLMEAPVKVIHRDLKS 107  
DB 1363 CIVTEWMGGSLRQFLTDHFNLLQONPHIRKLALDIAGKMYLHGWT-PILHRDLSSR 1421  
QY 108 NVVI-----AADGLKICDGSARFHNHTHM--SLVGTFPMAPEVIQSLP 152  
DB 1422 NILDHNIDPKNPVSSRODIKCKISDFGLSLKKEQASQMTQSVGCIPLYAPEVFKGDS 1481  
QY 153 VSETCDTYSYGVVLEMLTREVPFKGLEGLQVAVLVEKNERLTIPSSCPSPFAELLHOC 212  
DB 1482 NSEKSDVSYGMVLFELLTSDPEQDMKMKMAHAAAYESYRPPILPTSSKWEILTQC 1541  
QY 213 WEADAKRPSFKQIISILESM-----SNDTSLP 240  
DB 1542 WDSNPDSRPTFKQIIVHLKEMEDQGVSSFAVP 1574

RESULT 15

ABL\_FSVHY STANDARD; PRT; 439 AA.  
ID ABL\_FSVHY  
AC P10447;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Tyrosine-protein kinase transforming protein ABL (EC 2.7.1.112).  
GN V-ABL.  
OS Feline sarcoma virus (strain Hardy-Zuckerman 2).  
OS Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.  
OX NCBI\_TaxID=11776;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87141338; PubMed=3029415;  
RA Bergold P.J., Blumenthal J.A., D'Andrea E., Snyder H.W. Jr.,  
RA Lederman L., Silvestrone A., Nguyen H., Besmer P.;  
RT "Nucleic acid sequence and oncogenic properties of the H22 feline  
RT sarcoma virus v-abl insert.";  
RL J. Virol. 61:1193-1202(1987).  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL-POL  
CC POLYPROTEIN.  
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
CC SUBFAMILY.  
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M15805; AAA43042.1; -  
DR PIR; A26132; TVMVAB.  
DR HSSP; P00519; 1BBZ.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00017; SH2; 1.  
DR Pfam; PF00018; SH3; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00401; SH2DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.

DR ProDom; PD000001; Euk\_pkinase; 1.  
DR ProDom; PD000066; SH3; 1.  
DR ProDom; PD000093; SH2; 1.  
DR SMART; SM00252; SH2; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS50001; SH2; 1.  
DR PROSITE; PS50002; SH3; 1.  
KW Polypeptide; Tyrosine-protein kinase; Transferase; Oncogene;  
KW SH2 domain; SH3 domain.  
FT DOMAIN 10 70 SH3.  
FT DOMAIN 76 166 SH2.  
FT DOMAIN 191 439 PROTEIN KINASE.  
SQ SEQUENCE 439 AA; 50004 MW; 13579EDFED1481AB CRC64;

Query Match 27.0%; Score 351.5; DB 1; Length 439;  
Best Local Similarity 32.9%; Pred. No. 3.6e-21;  
Matches 78; Conservative 51; Mismatches 93; Indels 15; Gaps 7;

QY 1 YRAKWSIQDKEVAVKL-----LKIE---KEAELSVLSHRNITQFYGVILEPPNYGIYTE 53  
DB 206 YEGVWKYSLTVAVKTLKEDTMEVEEFLKEAAVMKEIKHPNLVQLGCVCTREPPFYITE 265  
QY 54 YASLSLYDYINSNRSEEMDMDHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVIAA 113  
DB 266 FMTYGNLDYLRECNROEVNAVLLYMATQISSAMEYLEKK---NFIHRDLAARNCLVGE 322  
QY 114 DGVLKICDFGASRFHNHTHMSLVGT-FP--WMAPEVIOQLPVSETCDTYSYGVLWEML 170  
DB 323 NHLVKVADFGISRLMTGDTYTAHAGTKPEPIKWTAPESLAYNKFISKSDVWAFGVLLWEIA 382  
QY 171 TRGV-PEKGLEGLQVAVLVEKNERLTIPSSCPRSFAELLHOCWEADAKRPSFKQI 226  
DB 383 TYGMSPYFGIDLSQV-YELLEKDYRMERPEGCEPEKYELMRACQWNPSPDRPAFAEI 438

Search completed: May 1, 2003, 20:36:20  
Job time : 15.4017 secs.



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:28:47 ; Search time 27.5593 Seconds

(without alignments)  
1846.697 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_31\_277

Perfect score: 1300

Sequence: 1 YRAKWSQDKAVAKKLKI.....SILESMSNDTSLPDKNSFL 247

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL 21: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*  
15: sp\_virus: \*  
16: sp\_bacteriophage: \*  
17: sp\_archaea: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description        |
|------------|-------|-------------|--------|-----------|--------------------|
| 1          | 1300  | 100.0       | 455    | 4 Q9HCC4  | Q9HCC4 homo sapien |
| 2          | 1300  | 100.0       | 800    | 4 Q9HDD2  | Q9HDD2 homo sapien |
| 3          | 1300  | 100.0       | 800    | 4 Q9HCC5  | Q9HCC5 homo sapien |
| 4          | 1300  | 100.0       | 800    | 4 Q9NYL2  | Q9NYL2 homo sapien |
| 5          | 1300  | 100.0       | 800    | 4 Q9NYE9  | Q9NYE9 homo sapien |
| 6          | 1289  | 99.2        | 454    | 11 Q9ESL3 | Q9ESL3 mus musculu |
| 7          | 1289  | 99.2        | 802    | 11 Q9ESL4 | Q9ESL4 mus musculu |
| 8          | 1064  | 81.8        | 371    | 13 Q9OZY8 | Q9OZY8 brachydanio |
| 9          | 511.5 | 39.3        | 406    | 10 Q23719 | Q23719 arabidopsis |
| 10         | 511.5 | 39.3        | 880    | 4 Q8S9K4  | Q8S9K4 arabidopsis |
| 11         | 508.5 | 39.1        | 1066   | 4 Q9H2N5  | Q9H2N5 homo sapien |
| 12         | 495.5 | 38.1        | 740    | 5 Q21982  | Q21982 caenorhabdi |
| 13         | 494   | 38.0        | 1001   | 11 Q8VDG6 | Q8VDG6 mus musculu |
| 14         | 490   | 37.7        | 859    | 4 Q8WY25  | Q8WY25 homo sapien |
| 15         | 483   | 37.2        | 564    | 4 Q9H1Y7  | Q9H1Y7 homo sapien |
| 16         | 482.5 | 37.1        | 847    | 4 Q16584  | Q16584 homo sapien |

|    |       |      |      |           |                    |
|----|-------|------|------|-----------|--------------------|
| 17 | 481.5 | 37.0 | 850  | 11 Q9JJI5 | Q9JJI5 mus musculu |
| 18 | 480   | 36.9 | 886  | 10 Q9LYI8 | Q9LYI8 arabidopsis |
| 19 | 475.5 | 36.6 | 1030 | 10 Q9C9U5 | Q9C9U5 arabidopsis |
| 20 | 474   | 36.5 | 570  | 4 Q8WMN2  | Q8WMN2 homo sapien |
| 21 | 474   | 36.5 | 1036 | 4 Q8WMN1  | Q8WMN1 homo sapien |
| 22 | 472.5 | 36.3 | 977  | 5 Q9VW24  | Q9VW24 drosophila  |
| 23 | 469.5 | 36.1 | 966  | 4 Q43283  | Q43283 homo sapien |
| 24 | 460.5 | 35.4 | 462  | 10 Q39886 | Q39886 glycine max |
| 25 | 453   | 34.8 | 982  | 10 Q65833 | Q65833 lycopersico |
| 26 | 452   | 34.8 | 903  | 10 Q9FPR5 | Q9FPR5 oryza sativ |
| 27 | 448   | 34.5 | 412  | 10 Q9M085 | Q9M085 arabidopsis |
| 28 | 445.5 | 34.3 | 1020 | 5 Q9W3I3  | Q9W3I3 drosophila  |
| 29 | 445.5 | 34.3 | 1148 | 5 Q95VF6  | Q95VF6 drosophila  |
| 30 | 445.5 | 34.3 | 1161 | 5 Q95UN8  | Q95UN8 drosophila  |
| 31 | 441   | 33.9 | 525  | 10 Q9FIL6 | Q9FIL6 arabidopsis |
| 32 | 440   | 33.8 | 982  | 10 Q93YG8 | Q93YG8 lycopersico |
| 33 | 438   | 33.7 | 525  | 10 Q8W022 | Q8W022 arabidopsis |
| 34 | 436.5 | 33.6 | 933  | 10 Q9FPR3 | Q9FPR3 arabidopsis |
| 35 | 435.5 | 33.5 | 847  | 10 Q93XL9 | Q93XL9 rosa hybrid |
| 36 | 434   | 33.4 | 855  | 5 Q01700  | Q01700 caenorhabdi |
| 37 | 432   | 33.2 | 957  | 10 Q9FPR4 | Q9FPR4 hordeum vul |
| 38 | 431.5 | 33.2 | 553  | 10 Q81808 | Q81808 arabidopsis |
| 39 | 431.5 | 33.2 | 570  | 10 Q8RWL6 | Q8RWL6 arabidopsis |
| 40 | 431   | 33.2 | 829  | 10 Q24027 | Q24027 lycopersico |
| 41 | 427   | 32.8 | 475  | 10 Q9STG5 | Q9STG5 arabidopsis |
| 42 | 427   | 32.8 | 637  | 10 Q94J41 | Q94J41 oryza sativ |
| 43 | 426   | 32.8 | 480  | 10 Q9FDV7 | Q9FDV7 lagus sylva |
| 44 | 422.5 | 32.5 | 2651 | 10 Q9FRR5 | Q9FRR5 arabidopsis |
| 45 | 421   | 32.4 | 806  | 10 Q9ZSD8 | Q9ZSD8 lycopersico |

## ALIGNMENTS

RESULT 1  
Q9HCC4 PRELIMINARY; PRT; 455 AA.  
ID Q9HCC4 AC Q9HCC4;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE MLTK-beta (Similar to sterile-alpha motif and leucine zipper  
DE containing kinase AZK) (Mixed lineage kinase) (Mixed lineage  
DE kinase-related kinase MRK-beta).  
GN MLTK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gotoh I., Adachi M., Nishida E.;  
RT "Identification and Characterization of a Novel MAP Kinase Kinase  
RT Kinase, MLTK.";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=COLON;  
RA Strausberg R.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Acton S.;  
RT "MLK-mixed lineage kinase.";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21950776; PubMed=11836244;  
RA Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;  
RT "MRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in  
RT gamma-Radiation-induced Cell Cycle Arrest.";  
RL J. Biol. Chem. 277:13873-13882(2002).  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.



DR EMBL; AB049734; BAB16445.1; -  
 DR EMBL; BC001401; AAH01401.1; -  
 DR EMBL; AF325454; AAK11615.1; -  
 DR EMBL; AF480462; AAL85892.1; -  
 DR HSSP; P12931; 1FMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TYRKc; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 455 AA; 51582 MW; E87DB84A4D58B752 CRC64;

Query Match 100.0%; Score 1300; DB 4; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-114;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWTISODKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILLEPPNYGIVTEYASLGSL 60  
 |||||||  
 Db 31 YRAKWTISODKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILLEPPNYGIVTEYASLGSL 90  
 QY 61 YDYINSNRSEEMDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLKIC 120  
 |||||||  
 Db 91 YDYINSNRSEEMDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLKIC 150  
 QY 121 DFGASRFHNHTTHMSLVGTFPWWAPVYIQLPVSETCDTYSYGVILWEMLTREVPFKGLE 180  
 |||||||  
 Db 151 DFGASRFHNHTTHMSLVGTFPWWAPVYIQLPVSETCDTYSYGVILWEMLTREVPFKGLE 210  
 QY 181 GLQVAVLVVEKNERLTIPSSCPSPFAELLHQWEADAKKRPSPFKQIISILEMSNDTSLP 240  
 |||||||  
 Db 211 GLQVAVLVVEKNERLTIPSSCPSPFAELLHQWEADAKKRPSPFKQIISILEMSNDTSLP 270  
 QY 241 DKCNSFL 247  
 |||||||  
 Db 271 DKCNSFL 277

## RESULT 2

Q9HDD2 PRELIMINARY; PRT; 800 AA.  
 Q9HDD2;  
 01-MAR-2001 (Tremblrel. 16, Created)  
 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Placible mixed-lineage kinase protein.  
 GN MLTKAK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LYMPHOID ORGAN;  
 RA Abe Y., Ueda N.;  
 RT "Placible Mixed-lineage kinase derived from LAK cell."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AB030034; BAB12040.1; -  
 DR HSSP; P12931; 1FMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.

DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TYRKc; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91155 MW; B2814509EC54B07A CRC64;

Query Match 100.0%; Score 1300; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-114;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWTISODKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILLEPPNYGIVTEYASLGSL 60  
 |||||||  
 Db 31 YRAKWTISODKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILLEPPNYGIVTEYASLGSL 90  
 QY 61 YDYINSNRSEEMDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLKIC 120  
 |||||||  
 Db 91 YDYINSNRSEEMDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLKIC 150  
 QY 121 DFGASRFHNHTTHMSLVGTFPWWAPVYIQLPVSETCDTYSYGVILWEMLTREVPFKGLE 180  
 |||||||  
 Db 151 DFGASRFHNHTTHMSLVGTFPWWAPVYIQLPVSETCDTYSYGVILWEMLTREVPFKGLE 210  
 QY 181 GLQVAVLVVEKNERLTIPSSCPSPFAELLHQWEADAKKRPSPFKQIISILEMSNDTSLP 240  
 |||||||  
 Db 211 GLQVAVLVVEKNERLTIPSSCPSPFAELLHQWEADAKKRPSPFKQIISILEMSNDTSLP 270  
 QY 241 DKCNSFL 247  
 |||||||  
 Db 271 DKCNSFL 277

## RESULT 3

Q9HCC5 PRELIMINARY; PRT; 800 AA.  
 Q9HCC5;  
 01-MAR-2001 (Tremblrel. 16, Created)  
 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE MLTK-alpha.  
 GN MLTK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21264927; PubMed=11042189;  
 RX Gotoh I., Adachi M., Nishida E.;  
 RA "Identification and Characterization of a Novel MAP Kinase Kinase  
 RT Kinase, MLTK."  
 RL J. Biol. Chem. 276:4276-4286(2001).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AB049733; BAB16444.1; -  
 DR HSSP; P12931; 1FMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TYRKc; 1.

DR PROSITE; PS50011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91188 MW; 2C8593824AB3FADD CRC64;

Query Match 100.0%; Score 1300; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-114;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSQDKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIIVTEYASLGSL 60  
 DB 31 YRAKWSQDKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIIVTEYASLGSL 90  
 QY 61 YDYINSRSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVIAADGVLKIC 120  
 DB 91 YDYINSRSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVIAADGVLKIC 150  
 QY 121 DFGASRFHNHTTHMSLVGTFPPWMAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLE 180  
 DB 151 DFGASRFHNHTTHMSLVGTFPPWMAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLE 210  
 QY 181 GLQVAVLVEKNERLTIPSSCPSPFAELLHQCEADAKKRPSFKQIISLESMSNDTSLP 240  
 DB 211 GLQVAVLVEKNERLTIPSSCPSPFAELLHQCEADAKKRPSFKQIISLESMSNDTSLP 270  
 QY 241 DKCNSFL 247  
 DB 271 DKCNSFL 277

## RESULT 4

Q9NYL2 PRELIMINARY; PRT; 800 AA.  
 AC Q9NYL2;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Mixed lineage kinase.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20384179; Pubmed=10924358;  
 RA Liu T.C., Huang C.J., Chu Y.C., Wei C.C., Chou C.C., Chou M.Y.,  
 RA Chou C.K., Yang J.J.;  
 RT "Cloning and expression of ZAK, a mixed lineage kinase-like protein  
 RT containing a leucine-zipper and a sterile-alpha motif.";  
 RT Biochem. Biophys. Res. Commun. 274:811-816(2000).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AF238255; AAF63490.1; -.  
 DR HSSP; P12931; 1FMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYKC; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91264 MW; DA82D7ABB2082F43 CRC64;

Query Match 100.0%; Score 1300; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-114;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSQDKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIIVTEYASLGSL 60  
 DB 31 YRAKWSQDKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIIVTEYASLGSL 90  
 QY 61 YDYINSRSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVIAADGVLKIC 120  
 DB 91 YDYINSRSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVIAADGVLKIC 150  
 QY 121 DFGASRFHNHTTHMSLVGTFPPWMAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLE 180  
 DB 151 DFGASRFHNHTTHMSLVGTFPPWMAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLE 210  
 QY 181 GLQVAVLVEKNERLTIPSSCPSPFAELLHQCEADAKKRPSFKQIISLESMSNDTSLP 240  
 DB 211 GLQVAVLVEKNERLTIPSSCPSPFAELLHQCEADAKKRPSFKQIISLESMSNDTSLP 270  
 QY 241 DKCNSFL 247  
 DB 271 DKCNSFL 277

## RESULT 5

Q9NYE9 PRELIMINARY; PRT; 800 AA.  
 AC Q9NYE9;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Sterile-alpha motif and leucine zipper containing kinase AZK (Mixed  
 DE lineage kinase-related kinase MRK-alpha).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA McNeel J.J., Prima N., Diamond T.E., Dower S.K., Guesdon F.;  
 RT "Cloning and characterisation of AZK, a mixed lineage kinase  
 RT containing a sterile-alpha motif.";  
 RT Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21950776; Pubmed=11836244;  
 RA Gross E.A., Cailow M.G., Waldbaum L., Thomas S., Ruggieri R.;  
 RA "MRK, a Mixed lineage Kinase-related Molecule That Plays a Role in  
 RT gamma-Radiation-induced Cell Cycle Arrest.";  
 RT J. Biol. Chem. 277:13873-13882(2002).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AF251441; AAF65822.1; -.  
 DR EMBL; AF480461; AAL85891.1; -.  
 DR HSSP; P12931; 1FMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYKC; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91181 MW; B289D836EC52B295 CRC64;

Query Match 100.0%; Score 1300; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-114;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YRAKWSQDKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIIVTEYASLGSL 60  
 |||

```

Db      31 YRAKWSIQDKEVAVKKLKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90
QY      61 YDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120
Db      91 YDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150
QY      121 DFGASRFHNHTTHMSLVGTFPPMAPEVIOQLPVSEICDTYSYGVVLWMLTREVPFKGLE 180
Db      151 DFGASRFHNHTTHMSLVGTFPPMAPEVIOQLPVSEICDTYSYGVVLWMLTREVPFKGLE 210
QY      181 GLQVAMLVEKNERLTIPSSCPSPFAELLHQCWEADAKKRPSEFKOIIISLESMSNDTSLP 240
Db      211 GLQVAMLVEKNERLTIPSSCPSPFAELLHQCWEADAKKRPSEFKOIIISLESMSNDTSLP 270
QY      241 DKCNSFL 247
Db      271 DKCNSFL 277

```

RESULT 6  
ID Q9ESL3 PRELIMINARY; PRT; 454 AA.

```

AC Q9ESL3;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE MLTK-beta.
GN ZAK OR MLTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264927; Pubmed=11042189;
RA Gotoh I., Adachi M., Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK.";
RL J. Biol. Chem. 276:4276-4286(2001).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049732; BAB16443.1; -.
DR HSSP; P12931; 1FMK.
DR MGD; MGI:1931274; Zak.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 454 AA; 51366 MW; 35C2FC0D729D9395 CRC64;

```

Query Match 99.2%; Score 1289; DB 11; Length 454;  
Best Local Similarity 98.8%; Pred. No. 1.6e-113;  
Matches 244; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 YRAKWSIQDKEVAVKKLKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60
Db      31 YRAKWSIQDKEVAVKKLKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90
QY      61 YDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120
Db      91 YDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150
QY      121 DFGASRFHNHTTHMSLVGTFPPMAPEVIOQLPVSEICDTYSYGVVLWMLTREVPFKGLE 180
Db      151 DFGASRFHNHTTHMSLVGTFPPMAPEVIOQLPVSEICDTYSYGVVLWMLTREVPFKGLE 210

```

```

QY      181 GLQVAMLVEKNERLTIPSSCPSPFAELLHQCWEADAKKRPSEFKOIIISLESMSNDTSLP 240
Db      211 GLQVAMLVEKNERLTIPSSCPSPFAELLHQCWEADAKKRPSEFKOIIISLESMSNDTSLP 270
QY      241 DKCNSFL 247
Db      271 DKCNSFL 277

```

RESULT 7  
ID Q9ESL4 PRELIMINARY; PRT; 802 AA.

```

AC Q9ESL4;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE MLTK alpha.
GN ZAK OR MLTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264927; Pubmed=11042189;
RA Gotoh I., Adachi M., Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK.";
RL J. Biol. Chem. 276:4276-4286(2001).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049731; BAB16442.1; -.
DR HSSP; P12931; 1FMK.
DR MGD; MGI:1931274; Zak.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00221; STYKc; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 802 AA; 91719 MW; D431DF8F312A43CC CRC64;

```

Query Match 99.2%; Score 1289; DB 11; Length 802;  
Best Local Similarity 98.8%; Pred. No. 3.4e-113;  
Matches 244; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 YRAKWSIQDKEVAVKKLKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60
Db      31 YRAKWSIQDKEVAVKKLKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90
QY      61 YDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120
Db      91 YDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150
QY      121 DFGASRFHNHTTHMSLVGTFPPMAPEVIOQLPVSEICDTYSYGVVLWMLTREVPFKGLE 180
Db      151 DFGASRFHNHTTHMSLVGTFPPMAPEVIOQLPVSEICDTYSYGVVLWMLTREVPFKGLE 210
QY      181 GLQVAMLVEKNERLTIPSSCPSPFAELLHQCWEADAKKRPSEFKOIIISLESMSNDTSLP 240
Db      211 GLQVAMLVEKNERLTIPSSCPSPFAELLHQCWEADAKKRPSEFKOIIISLESMSNDTSLP 270

```





RA Bowser L., Carninci P., Chang B., Dale J.M., Goldsmith A.D.,  
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.,  
 RT "Arabidopsis cDNA clones."  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY075641; AAL77650.1; -  
 SQ SEQUENCE 880 AA; 97881 MW; 5D9AD50C2B08444A CRC64;

Query Match 39.3%; Score 511.5; DB 10; Length 880;  
 Best Local Similarity 40.5%; Pred. No. 1.4e-39;  
 Matches 104; Conservative 51; Mismatches 85; Indels 17; Gaps 6;

QY 1 YRAKWSIQDEKAVAKKL-----KIEKAEILSVLSHRNIIQFYGVILEPPNYGI 50  
 DB 624 YRAEW--NGTEVAVKKFLDQDFSGDALTFKSEIEIMLRHPNVVLFMGAVTRPPNFSI 681  
 51 VTEYASLSLYDYINSNRSEEMDMHIMTWATDVAKGMHYLMEAPVKVIHRDLKSRNV 110  
 DB 682 LTFELPRGSLYRLH-RPNHQIDEXRRMALDVAKGMNYLHTSPT-VVHRDLKSPNLL 739  
 111 IAADGVLCICDFGASRFHNHT--THMSLVGTFPMAPEVIOSLPVSETCDTYSYGVLM 168  
 DB 740 VDKNMVVKVCDPGLSRMKHTYLSKSTAGTPEWMAPEVLNRPANKEKCDVYSGVLM 799  
 QY 169 MLTREVPEFKGLEQVAVLVEKNERLTIPSSCPSPFAELLHQCWEADAKRPSFKQIIS 228  
 DB 800 LATSRRVPMKGLNPMQVGVAGVFQNRLEIPDDIDLTVAAQIIRECWQTEPHLRPSFTQLMQ 859  
 QY 229 ILESMSNDTSLPDKCNS 245  
 DB 860 SLKRLQG-LNISNRANT 875

## RESULT 11

Q9H2N5 PRELIMINARY; PRT; 1066 AA.  
 AC Q9H2N5;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Mixed lineage kinase MLK1 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 SEQUENCE FROM N.A.  
 RA McNeer J.J., Dower S.K., Guesdon F.;  
 RT "cDNA sequence and gene organisation of mixed lineage kinase 1.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMBL; AF251442; AAG44591.1; -  
 DR HSSP; P29355; 1SEM.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.

DR PROSITE; PS00002; SH3; 1.  
 KW Kinase; SH3 domain.  
 FT NON\_TER 1  
 SQ SEQUENCE 1066 AA; 118463 MW; EDD08EBBE7482723 CRC64;  
 Query Match 39.1%; Score 508.5; DB 4; Length 1066;  
 Best Local Similarity 42.7%; Pred. No. 3.4e-39;  
 Matches 108; Conservative 43; Mismatches 77; Indels 25; Gaps 4;

QY 1 YRAKWSIQDEKAVK-----KLKIEKAEILSVLSHRNIIQFYGVILEPPNY 48  
 DB 107 YRAFWIGD--EAVKARHPDEDISQTIENVRQEAFLKHPNIIALRGVCLKEPNL 164  
 QY 49 GIVTEYASLSLYDYINSNRSEEMDMHIMTWATDVAKGMHYLMEAPVKVIHRDLKSRN 108  
 DB 165 CLVMEFARGGPLNRLVLSGR---IPDILVNWAVQIARGMNYLHDEAIVPIHRDLKSSN 221  
 QY 109 VVI-----AADGVLCICDFGASRFHNHTTHMSLVGTFPMAPEVIOSLPVSETCDTY 160  
 DB 222 ILLOKVENGDLSNKLKITDPGLAREWHRTTKMSAGTYAMMAPEVIRASMFSGSDVW 281  
 QY 161 SYGVVLWEMLTREVPEFKGLEQVAVLVEKNERLTIPSSCPSPFAELLHQCWEADAKR 220  
 DB 282 SYGVLLWELLTGEVFPFRGIDGLAVAYGVAMNKALPIPTCEPPEFAKLMECDWNPDSHR 341  
 QY 221 PSFKQIISLES 233  
 DB 342 PSFTNILDQLTTI 354

## RESULT 12

Q21982 PRELIMINARY; PRT; 740 AA.  
 AC Q21982;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Hypothetical protein R13F6.6.  
 GN R13F6.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 [1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 [2]  
 SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Miller N.;  
 RT "The sequence of C. elegans cosmid R13F6.";  
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.  
 [3]  
 SEQUENCE FROM N.A.  
 RP STRAIN=BRISTOL N2;  
 RC Waterston R.;  
 RT "Direct Submission."  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; U00046; AAC47047.3; -  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.

DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00219; TYRKC; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Hypothetical protein; Serine/threonine-protein kinase;  
Transferrase.  
SQ SEQUENCE 740 AA; 84075 MW; A36B698D1E1D3A0F CRC64;

Query Match 38.1%; Score 495.5; DB 5; Length 740;  
Best Local Similarity 41.0%; Pred. No. 3.5e-38;  
Matches 102; Conservative 48; Mismatches 86; Indels 13; Gaps 5;

QY 1 YRAKMSIQD---KEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPN-YGIVTEYAS 56  
Db 67 FSGNWTLPDGSQRTIALKKVFLKEAEILSKIRHKNIIQFYGICKATGNDFFIVTEYAE 126  
QY 57 LGSLYDYINSNRSEEM-----DMDHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNV 109  
Db 127 KGSLYDFIHSEESQSFASSSGNSFDVVVKWASQIASGIQYLHYDAVDTIHRDLKSKNV 186  
QY 110 VIAADGVLCIDFGASRFHNHT-THMSLVGFPPMAPE-VIQLPVSETCDTYSYGVLTW 167  
Db 187 VLDKMLVCKICDFTGSKDLTHSCTAPSWGTAAMSPEMILQSEGLTATDVWSYGVLTW 246  
QY 168 EMLTREVFPFKGLEGLQVAMLVKERNLTISSCPSPFAELHQCWEADAKRPSFKQII 227  
Db 247 ELJSKEVPYKDYSEFRIFTMTQSGITLAIIPSCPAFLKQLMSCWMTPKDRANRQIQ 306  
QY 228 SILESMSND 236  
Db 307 GELNRLAGN 315

## RESULT 13

Q8VDG6 PRELIMINARY; PRT; 1001 AA.  
ID Q8VDG6  
AC Q8VDG6;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Similar to mitogen-activated protein kinase kinase 9.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
P SEQUENCE FROM N.A.  
A Strauberg R.;

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
EMBL; BC021891; AAH21891.1;

DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00219; TYRKC; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
DR PROSITE; PS50002; SH3; 1.  
KW Kinase.

SQ SEQUENCE 1001 AA; 109983 MW; E10042C868B9953C CRC64;

Query Match 38.0%; Score 494; DB 11; Length 1001;  
Best Local Similarity 39.4%; Pred. No. 7.3e-38;  
Matches 108; Conservative 47; Mismatches 85; Indels 34; Gaps 5;

QY 1 YRAKMSIQDKEVAVKLLK-----IEKEAELISVLSHRNIIQFYGVILEPPNY 48  
Db 125 YRAIW--QGQEVAVKAAARDEQDAAAASVRREARLFAMLRHPNIIQLRGVCLRQPHL 182  
QY 49 GIVTEYASLGSLYDYI-----NSNRSEMDHMTWATDVAKGMHYLHMEAPVK 98  
Db 183 CLVEFARGGALNRLAALAAASDPAPGPRARRIPQVLVNMVAQIARGMLYLHEAVVP 242  
QY 99 VIHRDLKSRNVV-----AADGVLCIDFGASRFHNHTTHMSLVGFPPMAPEVIQS 150  
Db 243 ILHRDLKSNILLLEKIEHDDICNKLKITDFGLAREWHRTTRMSAGTYAMAPVIR 302  
QY 151 LPVSETCDTYSYGVLTWMLTREVFPFKGLEGLQVAMLVKERNLTISSCPSPFAELH 210  
Db 303 SLFSKGSIDWSYGVLTWMLTGEVPRYRGIDGLAVAYGAVVNKLTLPISGCEPFPALMK 362  
QY 211 QCWEADAKRPSFKQIISILESMSND--TSLPDK 242  
Db 363 ECWEQDPHIRPSFALILQOLTAIEEAVLTNMPQE 396

## RESULT 14

Q8WY25 PRELIMINARY; PRT; 859 AA.  
ID Q8WY25  
AC Q8WY25;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Zipper protein kinase.  
GN MAP3K12.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
P SEQUENCE FROM N.A.  
RA Reddy U.R.;  
RT "Genomic Structure and Promoter Characterization of the Homo sapiens  
MAP3K12 Gene.";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
EMBL; AF283475; AAL67158.1;

DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00219; TYRKC; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
KW Kinase.  
SQ SEQUENCE 859 AA; 93219 MW; 1E1BCAD2F6DFCFE8 CRC64;

Query Match 37.7%; Score 490; DB 4; Length 859;  
Best Local Similarity 44.3%; Pred. No. 1.4e-37;  
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

QY 10 KEVAVKLLKIEKEAEL--LSVLSHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSN 67  
Db 147 EEVAVKVYRDL-KETDIKHLRLKHPNIIITFKGVCTQAPCYCLMEFCAQGQLYEVLRAG 205  
QY 68 RSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASR- 126  
Db 206 RPTVPSL--LVDWSMGIAAGNVLHLH--KIHRDLKSPNMLITYDDVVKISDFGTSKE 260  
QY 127 FHNHTTHMSLVGFPPMAPEVIQSIPVSETCDTYSYGVLTWMLTREVFPFKGLEGLQVAM 186  
Db 261 LSDSKTMSFAGTVAMAPVIRNEPVSEKVDIWSFGVLTWMLTGEIPYKVDSSAIIW 320  
QY 187 LVKERNLTISSCPSPFAELHQCWEADAKRPSFKQIISILESMSNDT-SLPDK 242



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:12:12 ; Search time 47.5397 Seconds

(without alignments)  
692.324 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_31\_277

Perfect score: 1300

Sequence: 1 YRAKWIISQDKKAVKLLKI.....SILESMSNDTSLPDKCNFL 247

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Actual number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

1: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS27gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDS27gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDS27gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDS27gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID        | Description            |
|------------|-------|-------------|--------|----|-----------|------------------------|
| 1          | 1300  | 100.0       | 455    | 21 | AAB18657  | A human regulator      |
| 2          | 1300  | 100.0       | 455    | 21 | AAV83278  | Human survival reg     |
| 3          | 1300  | 100.0       | 455    | 21 | AAV84321  | A human cardiovascular |
| 4          | 1300  | 100.0       | 473    | 22 | AAAM25322 | Human protein sequ     |
| 5          | 1300  | 100.0       | 800    | 22 | AAAB71957 | Human TGF-beta rec     |
| 6          | 1300  | 100.0       | 800    | 22 | AAAB65673 | Novel protein kina     |
| 7          | 1253  | 96.4        | 349    | 22 | AAAG75571 | Human colon cancer     |
| 8          | 744   | 57.2        | 141    | 21 | AAAG03583 | Human secreted pro     |
| 9          | 497.5 | 38.3        | 1046   | 22 | AAE11775  | Human kinase (PKIN     |
| 10         | 497.5 | 38.3        | 1097   | 23 | AAE21717  | Human PKIN-12 prot     |

|    |       |      |      |    |           |                    |
|----|-------|------|------|----|-----------|--------------------|
| 11 | 490   | 37.7 | 859  | 16 | AAR82886  | Human leucine zipp |
| 12 | 490   | 37.7 | 859  | 18 | AAW31227  | Human leucine-zipp |
| 13 | 489   | 37.6 | 888  | 23 | ABB57049  | Mouse ischaemic co |
| 14 | 487.5 | 37.5 | 1021 | 23 | ABP61000  | Novel human protei |
| 15 | 483   | 37.2 | 719  | 22 | AAAB85513 | Human protein kina |
| 16 | 483   | 37.2 | 1036 | 23 | ABB80923  | Novel human protei |
| 17 | 482.5 | 37.1 | 847  | 23 | AAE22763  | Human mitogen acti |
| 18 | 472.5 | 36.3 | 977  | 22 | ABB71694  | Drosophila melanog |
| 19 | 470   | 36.2 | 144  | 22 | ABG06092  | Novel human diagno |
| 20 | 469.5 | 36.1 | 1490 | 22 | ABG19123  | Novel human diagno |
| 21 | 453   | 34.8 | 982  | 22 | AAAB50439 | Tomato TCTR2. Lyc  |
| 22 | 452   | 34.8 | 903  | 22 | AAAB50440 | Rice EDRL. Oryza   |
| 23 | 448.5 | 34.5 | 850  | 23 | AAE18529  | Melon constitutive |
| 24 | 448   | 34.5 | 369  | 21 | AAAG22172 | Arabidopsis thalia |
| 25 | 448   | 34.5 | 374  | 21 | AAAG22171 | Arabidopsis thalia |
| 26 | 448   | 34.5 | 412  | 21 | AAAG22170 | Arabidopsis thalia |
| 27 | 445.5 | 34.3 | 1020 | 22 | ABB58999  | Drosophila melanog |
| 28 | 442   | 34.0 | 821  | 16 | AAR80574  | Arabidopsis CTPI p |
| 29 | 442   | 34.0 | 821  | 18 | AAW17938  | Constitutive tripl |
| 30 | 442   | 34.0 | 821  | 22 | AAAB50438 | Arabidopsis thalia |
| 31 | 438   | 33.7 | 933  | 22 | AAAB50437 | Arabidopsis thalia |
| 32 | 436.5 | 33.6 | 933  | 22 | AAAB50443 | Arabidopsis thalia |
| 33 | 435.5 | 33.5 | 970  | 22 | AAAB50443 | Arabidopsis thalia |
| 34 | 424   | 32.6 | 92   | 22 | AAU87295  | Barley EDRL. Hord  |
| 35 | 424   | 32.6 | 92   | 22 | AAU17234  | Novel signal trans |
| 36 | 418.5 | 32.2 | 589  | 21 | AAAG45984 | Arabidopsis thalia |
| 37 | 418.5 | 32.2 | 732  | 21 | AAAG45983 | Arabidopsis thalia |
| 38 | 418.5 | 32.2 | 760  | 21 | AAAG45982 | Arabidopsis thalia |
| 39 | 418   | 32.2 | 367  | 21 | AAAG32053 | Arabidopsis thalia |
| 40 | 418   | 32.2 | 369  | 21 | AAAG32052 | Arabidopsis thalia |
| 41 | 418   | 32.2 | 407  | 21 | AAAG32051 | Arabidopsis thalia |
| 42 | 399   | 30.7 | 678  | 22 | ABBS8061  | Drosophila melanog |
| 43 | 395   | 30.4 | 567  | 20 | AAV28998  | Human TGF-beta act |
| 44 | 395   | 30.4 | 579  | 18 | AAW27092  | Mouse transforming |
| 45 | 395   | 30.4 | 579  | 18 | AAW27093  | Human transforming |

ALIGNMENTS

|          |  |
|----------|--|
| RESULT 1 |  |
| AAB18657 |  |
| ID       | AAB18657 standard; Protein; 455 AA.                                    |
| AC       | AAB18657;  |
| XX       |  |
| DT       | 22-JAN-2001 (first entry)  |
| XX       |  |
| DE       | A human regulator of intracellular phosphorylation.                    |
| XX       |  |
| KW       | Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma; |
| KW       | neurological disorder; Parkinson's disease; demyelinating disease;     |
| KW       | meningitis; developmental disorder; neuromuscular disorder; cancer;    |
| KW       | myasthenia gravis; cell proliferative disorder; actinic keratosis;     |
| KW       | arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;    |
| KW       | autoimmune disorder; inflammatory disorder; Addison's disease;         |
| KW       | acquired immunodeficiency disease; allergy; diabetes mellitus;         |
| KW       | rheumatoid arthritis; microbial infection; trauma.                     |
| XX       |  |
| OS       | Homo sapiens.  |
| XX       |  |
| PH       | Key  |
| FT       | Domain   |
| FT       | Modified-site.   |
| FT       | Location/Qualifiers  |
| FT       | 16..257  |
| FT       | /note= "eukaryotic protein kinase domain"                              |
| FT       | 61   |
| FT       | /note= "potential phosphorylation site"                                |
| FT       | 89   |
| FT       | /note= "potential phosphorylation site"                                |
| FT       | 96   |
| FT       | /note= "potential phosphorylation site"                                |
| FT       | 97   |
| FT       | /note= "potential phosphorylation site"                                |
| FT       | 129..141   |
| FT       | Binding-site   |



|    |   |
|----|---|
| FT | /note= "protein kinase ATP-binding site"                                |
| PT | 159 /note= "potential glycosylation site"                               |
| FT | 234 /note= "potential phosphorylation site"                             |
| FT | 252 /note= "potential phosphorylation site"                             |
| FT | 258 /note= "potential phosphorylation site"                             |
| FT | 265 /note= "potential glycosylation site"                               |
| FT | 268 /note= "potential phosphorylation site"                             |
| FT | 294..322 /note= "leucine zipper"  |
| FT | 302 /note= "potential phosphorylation site"                             |
| FT | 302 /note= "potential phosphorylation site"                             |
| FT | 302 /note= "potential phosphorylation site"                             |
| FT | 342 /note= "potential phosphorylation site"                             |
| FT | 343 /note= "potential phosphorylation site"                             |
| FT | 346 /note= "potential phosphorylation site"                             |
| FT | 364 /note= "potential phosphorylation site"                             |
| FT | 409 /note= "potential glycosylation site"                               |
| FT | 410 /note= "potential phosphorylation site"                             |
| FT | 414 /note= "potential phosphorylation site"                             |
| FT | 415 /note= "potential phosphorylation site"                             |
| FT | 429 /note= "potential phosphorylation site"                             |
| FT | 434 /note= "potential phosphorylation site"                             |
| PN | WO200055332-A2.   |
| PD | 21-SEP-2000.  |
| PF | 17-MAR-2000; 2000WO-US07277.  |
| PR | 18-MAR-1999; 99US-0125593.  |
| PR | 20-MAY-1999; 99US-0135049.  |
| PR | 09-JUL-1999; 99US-0143188.  |
| XX | (INCY-) INCYTE PHARM INC.   |
| XX | Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;            |
| PI | Iu DAM, Au-Young J;   |
| XX | WPI: 2000-602121/57.  |
| DR | N-PSDB; AAA75674.   |
| XX | Novel human intracellular phosphorylation regulator polypeptides and    |
| PT | polynucleotides for diagnosis, prevention and treatment of              |
| PT | neurological, cell proliferative and autoimmune/inflammatory disorders  |
| PS | -   |
| XX | Claim 1; Page 75-76; 96pp; English.                                     |
| CC | The present sequence represents a human regulator of intracellular      |
| CC | phosphorylation (HRIP). HRIP is useful for screening agonists and       |
| CC | antagonists of HRIP polypeptide. HRIP and its agonist or antagonist     |
| CC | are useful for treating a disease or condition associated with          |
| CC | decreased or increased expression of functional HRIP. Diseases treated  |
| CC | or diagnosed include neurological disorders such as stroke, Parkinson's |
| CC | disease, demyelinating diseases, bacterial and viral meningitis and     |
| CC | other developmental disorders of the central nervous system,            |

neurovascular disorders, myasthenia gravis, cell proliferative disorders such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer including leukaemia, melanoma, myeloma and cancer of the adrenal gland, bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/inflammatory disorder such as Addison's disease, acquired immunodeficiency diseases, allergies, bronchitis, diabetes mellitus, rheumatoid arthritis, microbial infection and trauma.

Sequence 455 AA;

Query Match 100.0%; Score 1300; DB 21; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWISQDKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILLEPPNYGIVTEYASLSGL 60  
 |||||  
 DB 31 YRAKWISQDKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILLEPPNYGIVTEYASLSGL 90  
 |||||  
 QY 61 YDYINSNRSEEMDMDHIMTWA TDVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLKIC 120  
 |||||  
 DB 91 YDYINSNRSEEMDMDHIMTWA TDVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLKIC 150  
 |||||  
 QY 121 DFGASRFHNHTTHMSLVGTFPMNAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLE 180  
 |||||  
 DB 151 DFGASRFHNHTTHMSLVGTFPMNAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLE 210  
 |||||  
 QY 181 GLQVAVLVEKNERLTIPSSCPRFAELLHQCWEADAKKRPFKQIISILESMSNDTSLP 240  
 |||||  
 DB 211 GLQVAVLVEKNERLTIPSSCPRFAELLHQCWEADAKKRPFKQIISILESMSNDTSLP 270  
 |||||  
 QY 241 DKCNSFL 247  
 |||||  
 DB 271 DKCNSFL 277

RESULT 2  
 ID AAY83278 standard; Protein; 455 AA.  
 AAY83278  
 AC AAY83278;  
 XX  
 DT 16-AUG-2000 (first entry)  
 XX  
 DE Human survival regulating kinase (SRK).  
 KW Survival regulating kinase; SRK; Raf; MBP; BAD; MAPKK; MAPKK;  
 KW HAX-1; protein kinase; autophosphorylation; cell growth; regulation;  
 KW apoptosis; cell survival; nuclear targeting; tumour; human;  
 KW autoimmune disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200022142-A2.  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 20-SEP-1999; 99WO-US22008.  
 XX  
 PR 13-OCT-1998; 98US-0104088.  
 XX  
 PA (ONYX-) ONYX PHARM INC.  
 XX  
 PI Ruggieri R, Callow M, Diaz P;  
 DR MPI; 2000-317994/27.  
 DR N-PSDB; AAZ93783.  
 XX  
 PT Novel human survival regulating kinase polypeptide for screening agents  
 PT which modulate biological pathways associated with SRK useful in  
 PT treating autoimmune diseases, tumors and apoptosis-related disorders  
 PS Claim 4; Figure 2; 62pp; English.  
 XX

CC Survival regulating kinases (SRK) are a class of proteins involved in  
CC cell signal transduction pathways such as mitogen-activated protein  
CC kinase pathways. A protein kinase activity means that the SRK can  
CC catalyse a reaction in which a phosphate group is transferred from a  
CC phosphate donor to a phosphate acceptor amino acid residue,  
CC preferably the hydroxyl side chain of a serine or threonine.  
CC Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase  
CC activity is similar to that of a MAPKK such as Raf. has a range of  
CC other activities including a cell growth-regulatory activity, a cell  
CC survival promoting activity, a HAX-1 binding activity, an apoptosis  
CC suppressing activity a MAPK activation or stimulatory activity, a  
CC nuclear targeting activity and a SRK-specific immunogenic activity.  
CC SRK is useful for identifying agents which modulate cellular  
CC transformations mediated by Ras and SRK and agents that modulate the  
CC apoptosis suppression activity of SRK. This information may be useful  
CC in the treatment of autoimmune diseases, tumours and apoptosis  
CC related disorders.

XX Sequence 455 AA;

Query Match 100.0%; Score 1300; DB 21; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.9e-127;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSIQDKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLGS 60  
DB 31 YRAKWSIQDKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLGS 90  
QY 61 YDYINSNRSEEMDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120  
DB 91 YDYINSNRSEEMDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150  
QY 121 DFGASRFHNHTTHMSLVGTFPMWAPVYIQLPVSETCDTYSYGVLWEMLTREVPFKGLE 180  
DB 151 DFGASRFHNHTTHMSLVGTFPMWAPVYIQLPVSETCDTYSYGVLWEMLTREVPFKGLE 210  
QY 181 GLQVAVLVEKNERLTIPSSCPSPFAELLHQCEADAKKRPSPFKQIISILEMSNDTSLP 240  
DB 211 GLQVAVLVEKNERLTIPSSCPSPFAELLHQCEADAKKRPSPFKQIISILEMSNDTSLP 270  
QY 241 DKCNSFL 247  
DB 271 DKCNSFL 277

RESULT 3  
AAY84321

AAY84321 standard; Protein; 455 AA.

AAY84321;

DT 12-JUL-2000 (first entry)

DE A human cardiovascular system associated protein kinase-2.

KW Human; cardiovascular system associated protein kinase-2; CSAPK-2;  
KW signalling pathway; cell growth; cell differentiation; gene mapping;  
KW tissue typing; forensic identification; cardiovascular disease;  
KW congestive heart failure; transgenic animal.

OS Homo sapiens.

PN WO200014212-A1.

PD 16-MAR-2000.

PF 09-SEP-1999; 99WO-US20631.

PR 09-SEP-1998; 98US-0099657.

PR 29-SEP-1998; 98US-0163115.

PA (MILL-) MILLENNIUM PHARM INC.

PI Acton S;

XX WPI; 2000-271053/23.

DR N-PSDB; AAZ99726, AAZ99727.

XX

PT New nucleic acid encoding cardiovascular system associated protein  
PT kinase, used e.g. for diagnosis, treatment and prevention of  
PT cardiovascular disease

PS Claim 2; Fig 2; 163pp; English.

CC The present sequence represents a human cardiovascular system associated  
CC protein kinase-2 (CSAPK-2). CSAPK polypeptides are involved in signalling  
CC pathways associated with cell growth and differentiation. The CSAPK  
CC polypeptides and polynucleotides are used to screen for agents that  
CC specifically modulate CSAPK, which are potential therapeutic agents.  
CC They are also used for diagnosis, prognosis or monitoring of  
CC CSAPK-related diseases, gene mapping, tissue typing and forensic  
CC identification, and for treating or preventing disorders associated  
CC with aberrant CSAPK expression or activity, especially cardiovascular  
CC diseases such as congestive heart failure. They can also be used in  
CC pharmacogenomics. The CSAPK polynucleotide may also be used to generate  
CC transgenic animals.

XX Sequence 455 AA;

Query Match 100.0%; Score 1300; DB 21; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.9e-127;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSIQDKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLGS 60  
DB 31 YRAKWSIQDKEVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEYASLGS 90  
QY 61 YDYINSNRSEEMDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120  
DB 91 YDYINSNRSEEMDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150  
QY 121 DFGASRFHNHTTHMSLVGTFPMWAPVYIQLPVSETCDTYSYGVLWEMLTREVPFKGLE 180  
DB 151 DFGASRFHNHTTHMSLVGTFPMWAPVYIQLPVSETCDTYSYGVLWEMLTREVPFKGLE 210  
QY 181 GLQVAVLVEKNERLTIPSSCPSPFAELLHQCEADAKKRPSPFKQIISILEMSNDTSLP 240  
DB 211 GLQVAVLVEKNERLTIPSSCPSPFAELLHQCEADAKKRPSPFKQIISILEMSNDTSLP 270  
QY 241 DKCNSFL 247  
DB 271 DKCNSFL 277

RESULT 4

AAM25322

ID AAM25322 standard; Protein; 473 AA.

AC AAM25322;

DT 16-OCT-2001 (first entry)

DE Human protein sequence SEQ ID NO:837.

KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
KW antibacterial; endocrine; cardiac; central nervous system; virucide;  
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
KW dermatological; antiallergic; antidiabetic; osteopathic; eczema;  
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;

KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder.

OS Homo sapiens.

XX WO200153455-A2.

XX 26-JUL-2001.

XX 22-DEC-2000; 2000WO-US35017.

XX 23-DEC-1999; 99US-0471275.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457603/49.

XX N-PSDB; AAH99263.

XX Isolated human polynucleotides encoding polypeptides, useful for the

XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection -

XX Claim 20; Page 191; 1217pp; English.

XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
XX AAM25963. The proteins can have activities based on the tissues and  
XX cells they are expressed in, such as: antiinflammatory; antirheumatic;  
XX antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;  
XX central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
XX cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;  
XX antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;  
XX antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;  
XX antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
XX encoding them can be used in gene therapy, antisense therapy and vaccine  
XX production. The proteins and polynucleotides are useful for screening for  
XX agonists or antagonists of a protein and for the treatment and diagnosis  
XX of disorders associated with the activity of a protein e.g. inflammation,  
XX rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
XX neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
XX infections, autoimmunity, genetic diseases, haematopoietic disorders,  
XX anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
XX osteoporosis, severe combined immunodeficiency, eczema, allergic  
XX rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
XX Alzheimer's disease, Parkinson's disease, neurodegenerative and  
XX neurological disorders.

XX Sequence 473 AA;

XX Query Match 100.0%; Score 1300; DB 22; Length 473;

XX Best Local Similarity 100.0%; Pred. No. 3.1e-127; Indels 0; Gaps 0;

XX Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 YRAKWSQDKEVAVKLLKIEKAEILSVLSHRNIQFYGVILEPPNYGIVTEYASIGSL 60

XX 49 YRAKWSQDKEVAVKLLKIEKAEILSVLSHRNIQFYGVILEPPNYGIVTEYASIGSL 108

XX 61 YDIYNSNRSEMDMHTMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLKIC 120

XX 109 YDIYNSNRSEMDMHTMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLKIC 168

XX 121 DFGASRFHNHTTMSLVGTFPMAPEVIOQLPVSETCDTYSYGVVLWEMLTREVPFKGLE 180

XX 169 DFGASRFHNHTTMSLVGTFPMAPEVIOQLPVSETCDTYSYGVVLWEMLTREVPFKGLE 228

XX 181 GIQVAVLVEKNERLTIPSSCPSPFAELLHQCEWADAKKRPSPFOIISLESMSNDTSLP 240

XX 229 GIQVAVLVEKNERLTIPSSCPSPFAELLHQCEWADAKKRPSPFOIISLESMSNDTSLP 288

XX 241 DKCNSFL 247

DB 289 DKCNSFL 295

RESULT 5  
ID AAB71957 standard; Protein; 800 AA.

XX AAB71957;

XX 11-MAY-2001 (first entry)

XX Human TGF-beta receptor encoded by cDNA clone HDPSM48.

XX Human; antisclerotic; dermatological; immunosuppressive; cytostatic;  
XX antiinflammatory; anti-HIV; immunostimulant; cardiant; vascular;  
XX ophthalmological; neuroprotective; nootropic; anticonvulsant; vaccine;  
XX antiparkinsonian; antimicrobial; vulnery; gene therapy; infection;  
XX transforming growth factor; TGF; TGF-beta receptor; immune disorder;  
XX hyperproliferative disorder; cardiovascular disease; angiogenesis;  
XX neurological disorder.

XX Homo sapiens.

XX WO200112670-A1.

XX 22-FEB-2001.

XX 10-AUG-2000; 2000WO-US21736.

XX 13-AUG-1999; 99US-0148682.

XX 20-SEP-1999; 99US-0154887.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Ni J;

XX WPI; 2001-202858/20.

XX N-PSDB; AAF75336.

XX Nucleic acid molecules encoding 12 transforming growth factor-beta  
XX receptor polypeptides, useful for preventing, diagnosing and treating  
XX e.g. cancers, Parkinson's disease and diabetic retinopathy -  
XX Claim 11; Page 293-295; 311pp; English.

XX The present sequence is one of 12 novel human transforming growth factor  
XX (TGF)-beta receptor polypeptides. The TGF-beta receptor polynucleotides  
XX and polypeptides may be used in the prevention, diagnosis and treatment  
XX of diseases associated with inappropriate polypeptide expression. Such  
XX diseases include immune disorders (e.g. multiple sclerosis, systemic  
XX lupus erythematosus and human immuno-deficiency virus (HIV) infections),  
XX hyperproliferative disorders (e.g. cancers and Gaucher's disease),  
XX cardiovascular diseases (e.g. Scimitar syndrome, Chaga's cardiomyopathy  
XX and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft  
XX neovascularisation and diabetic retinopathy), neurological disorders  
XX (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease)  
XX and infectious diseases. The polynucleotides and polypeptides are also  
XX useful for promoting wound healing, regeneration and/or chemotaxis. The  
XX polynucleotides and their complementary sequences may also be used as DNA  
XX probes in diagnostic assays to detect and quantitate the presence of  
XX similar nucleic acid sequences in samples. The polypeptides may be used  
XX as antigens in the production of antibodies and in assays to identify  
XX modulators of protein expression and activity. The anti-TGF-beta receptor  
XX antibodies may be used to down regulate expression and activity and as  
XX diagnostic agents for detecting the presence of the polypeptides in  
XX samples.

XX Sequence 800 AA;

XX Query Match 100.0%; Score 1300; DB 22; Length 800;

XX Best Local Similarity 100.0%; Pred. No. 6.6e-127; Indels 0; Gaps 0;

XX Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWISQDKEVAVKKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEVASLGS 60  
Db 31 YRAKWISQDKEVAVKKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEVASLGS 90  
QY 61 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120  
Db 91 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150  
QY 121 DFGASRFHNHTHMSLVGTFPMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 180  
Db 151 DFGASRFHNHTHMSLVGTFPMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 210  
QY 181 GLOVAMLVVEKNERLTIPTSSCPSPFAELLHQCEADAKRPSFKQIISILESMSNDTSLP 240  
Db 211 GLOVAMLVVEKNERLTIPTSSCPSPFAELLHQCEADAKRPSFKQIISILESMSNDTSLP 270  
QY 241 DKCNSFL 247  
271 DKCNSFL 277

RESULT 6

AAB65673  
ID AAB65673 standard; Protein; 800 AA.

XX AAB65673;  
AC  
XX  
DT 27-MAR-2001 (first entry)

XX Novel protein kinase, SEQ ID NO: 201.

XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
KW immunosuppressive; cardiac; renal; antinflammatory; antiasthmatic;  
KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;  
KW immune disorder; cardiovascular disease; neurodegenerative disease;  
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.

OS Homo sapiens.

XX WO200073469-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14842.

XX 28-MAY-1999; 99US-0136503.

XX (SUGEN-) SUGEN INC.

PI Plowman GD, Martinez R, Whyte D, Sudersanam S;

DR WPI; 2001-032161/04.  
DR N-PSDB; AAF44701.

PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
PT treating immune-related diseases and disorders, cardiovascular disease,  
PT neurodegenerative diseases and/or cancers -

PS Claim 10; Fig 1; 310pp; English.

XX The present sequence is a novel protein kinase. The novel protein kinases  
CC and the nucleic acids that encode them may be used in the treatment and  
CC diagnosis of diseases associated with inappropriate kinase expression  
CC such as immune-related diseases and disorders, cardiovascular disease,  
CC neurodegenerative diseases and/or cancers. The nucleic acids and  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays. The kinase polypeptides may be used as antigens in the production  
CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
CC and kinase antagonists may also be used to down regulate kinase  
CC expression and activity. Diseases related to kinase expression and  
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune

CC disorders, complications of organ transplantation, myocardial infarction,  
CC immune disorders, cardiomyopathies, strokes, renal failure,  
CC oxidative-stress related disorders, chronic inflammatory bowel disease,  
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
CC reproductive disorders.

XX Sequence 800 AA;

Query Match 100.0%; Score 1300; DB 22; Length 800;  
Best Local Similarity 100.0%; Pred. NO. 6.6e-127;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWISQDKEVAVKKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEVASLGS 60  
Db 31 YRAKWISQDKEVAVKKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEVASLGS 90

QY 61 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120  
Db 91 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150

QY 121 DFGASRFHNHTHMSLVGTFPMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 180  
Db 151 DFGASRFHNHTHMSLVGTFPMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 210

QY 181 GLOVAMLVVEKNERLTIPTSSCPSPFAELLHQCEADAKRPSFKQIISILESMSNDTSLP 240  
Db 211 GLOVAMLVVEKNERLTIPTSSCPSPFAELLHQCEADAKRPSFKQIISILESMSNDTSLP 270

QY 241 DKCNSFL 247  
Db 271 DKCNSFL 277

RESULT 7

AAG75571  
ID AAG75571 standard; Protein; 349 AA.

AC AAG75571;

DT 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:6335.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.  
XX 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI; 2001-235357/24.  
DR N-PSDB; AAH34976.

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers -

PS Claim 11; Page 7789-7790; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon



CC cancer antigens have cytostatic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing  
CC inactive proteins or to supplement the patients own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated Ps;  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
CC and AAB77789 represent sequences used in the exemplification of the  
CC present invention.  
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 349 AA;

Query Match 96.4%; Score 1253; DB 22; Length 349;  
Best Local Similarity 100.0%; Pred. No. 1.6e-122;  
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRAKMWISODKEVAVKLLKLEKEAELISVLSHRNIIQFYGVILEPPNYGIATVEYASLGSL 60  
DB 87 YRAKMWISODKEVAVKLLKLEKEAELISVLSHRNIIQFYGVILEPPNYGIATVEYASLGSL 146  
OY 61 YDYINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLIKIC 120  
DB 147 YDYINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLIKIC 206  
OY 121 DFGASRFHNHTTHMSLVGTFPMAAPEVIOQLPVSETCDITYSYGVVLWEMLTREVPFKGLE 180  
DB 207 DFGASRFHNHTTHMSLVGTFPMAAPEVIOQLPVSETCDITYSYGVVLWEMLTREVPFKGLE 266  
OY 181 GIOVAVLVEKNERLTIPSSCPSPFAELLHQCEADAKKRPFKQIISILESMGNDTSL 239  
DB 267 GIOVAVLVEKNERLTIPSSCPSPFAELLHQCEADAKKRPFKQIISILESMGNDTSL 325

RESULT 8  
AAG03583  
ID AAG03583 standard; Protein; 141 AA.  
XX AAG03583;  
AC AAG03583;  
XX AAG03583;  
DT 06-OCT-2000 (first entry)

Human secreted protein, SEQ ID NO: 7664.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping.

OS Homo sapiens.  
XX EPI033401-A2.  
XX 06-SEP-2000.  
XX 21-FEB-2000; 2000EP-0200610.  
XX 26-FEB-1999; 99US-0122487.  
XX (GEST ) GENSET.  
XX Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
XX N-PSDB; AAC03589.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 13; SEQ ID 7664; 71pp + CD-ROM; English.

PS The present sequence is a polypeptide encoded by one of a large number  
XX of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
CC different tissues. EST sequences usually correspond mainly to the 3'  
CC untranslated region (UTR) of the mRNA because they are often obtained  
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
CC those cases where longer cDNA sequences have been obtained, the full 5'  
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
CC ends and can therefore be used to obtain full length cDNAs and genomic  
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
CC chromosome mapping procedures. They are used to obtain upstream  
CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 141 AA;

Query Match 57.2%; Score 744; DB 21; Length 141;  
Best Local Similarity 98.6%; Pred. No. 8.8e-70;  
Matches 139; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 72 MDMDHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLIKICDFGASRFHNHT 131  
DB 1 MDMDHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLIKICDFGASRFHNHT 60  
OY 132 THMSLVGTFPMAAPEVIOQLPVSETCDITYSYGVVLWEMLTREVPFKGLEGLQVAVLVVEK 191  
DB 61 THMSLVGTFPMAAPEVIOQLPVSETCDITYSYGVVLWEMLTREVPFKGLEGLQVAVLVVXX 120  
OY 192 NERLTIPSSCPSPFAELLHQCEADAKKRPFKQIISILESMGNDTSL 212  
DB 121 NERLTIPSSCPSPFAELLHQCEADAKKRPFKQIISILESMGNDTSL 141

RESULT 9  
AAE11775  
ID AAE11775 standard; Protein; 1046 AA.  
XX AAE11775;  
AC AAE11775;  
XX AAE11775;  
DT 18-DEC-2001 (first entry)

Human kinase (PKIN)-9 protein.

XX Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;  
KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;  
KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;  
KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;  
KW myasthenia gravis; cirrhosis; cataract; growth and development disorder;  
KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;  
KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;  
KW obesity; restorative therapy; immunomodulatory; vaccine; cardiovascular;  
KW antimicrobial; cytostatic; antiinflammatory; asthma.

XX Homo sapiens.

XX Location/Qualifiers  
XX 55..114  
FH Domain  
FT /note= "SH3 domain"  
FT 134..393  
FT /note= "Eukaryotic protein kinase domain"  
FT 136..386  
FT /note= "Protein kinase domain"  
FT 154..207  
FT /note= "Receptor tyrosine kinase"  
FT 181..228  
FT /note= "Receptor tyrosine kinase"  
FT 210..223  
FT /note= "Tyrosine kinase catalytic site"  
FT 232..254  
FT /note= "Receptor tyrosine kinase"

FT Region 248..266  
/note= "Tyrosine kinase catalytic site"  
FT Region 290..337  
/note= "Receptor tyrosine kinase"  
FT Region 291..340  
/note= "Receptor tyrosine kinase"  
FT Region 298..330  
/note= "Receptor tyrosine kinase"  
FT Region 301..311  
/note= "Tyrosine kinase catalytic site"  
FT Region 320..342  
/note= "Tyrosine kinase catalytic site"  
FT Region 337..389  
/note= "Receptor tyrosine kinase"  
FT Region 345..389  
/note= "Receptor tyrosine kinase"  
FT Region 356..404  
/note= "Receptor tyrosine kinase"  
FT Region 364..386  
/note= "Tyrosine kinase catalytic site"

MO200181555-A2.

01-NOV-2001.

20-APR-2001; 2001WO-US12992.

20-APR-2000; 2000US-199021P.  
28-APR-2000; 2000US-200226P.  
05-MAY-2000; 2000US-202339P.  
11-MAY-2000; 2000US-203505P.  
18-MAY-2000; 2000US-205564P.  
26-MAY-2000; 2000US-207739P.  
01-JUN-2000; 2000US-208795P.

(INCY-) INCYTE GENOMICS INC.

Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;  
Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Walla NK, Yao MG;  
Patterson C, Burrill JD, Marcus GA, Zingler KA, Recipon SA, Lu Y;  
Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;  
Walsh RT, Ramkumar J, Borowsky ML, Au-young J, Hillman JL;  
Gururajan R;

WPI; 2001-611740/70.  
N-PSDB; AAD18824.

Human kinases and nucleic acids, useful for preventing diagnosing and  
treating cancers, inflammation and immune disorders

Claim 1; Page 134-136; 166pp; English.

The present invention relates to human kinases (PKIN) and the nucleic  
acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is  
used in the prevention, diagnosis and treatment of diseases cancers,  
adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,  
acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,  
gout, microbial infections, cardiovascular disease and/or inflammation,  
myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial  
infarction, cataract, growth and development disorder, seizure disorder,  
pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage  
disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.  
PKIN may be used to treat disorders associated with decreased PKIN  
expression by rectifying mutations or deletions in a patient's genome  
that affect the activity of PKIN by expressing inactive proteins or to  
supplement the patient's own production of PKIN. PKIN nucleic acids may be  
used to produce the PKIN polypeptide, by inserting the nucleic acids into  
a host cell and culturing the cell to express the protein. PKIN nucleic  
acid and its complementary sequences may also be used as DNA probes in  
diagnostic assays to detect and quantitate the presence of similar  
nucleic acid sequences in samples and therefore which patients may be  
in need of restorative therapy. The present sequence is human PKIN-9  
protein.

XX SQ Sequence 1046 AA;  
Query Match 38.3%; Score 497.5; DB 22; Length 1046;  
Best Local Similarity 42.3%; Pred. No. 1.1e-42;  
Matches 107; Conservative 43; Mismatches 78; Indels 25; Gaps 4;  
QY 1 YRAKWSQDKEVAVK-----KLKIEKEAEILSVLSHRNITQFYGVILEPPNY 48  
Db 149 YRAFWIGD--EVAVKARHDPDEDISQTIENVROEAKLFAMLKHPNIIALRGVCLKEPNL 206  
QY 49 GIVTEYASLSLYDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRN 108  
Db 207 CLVMEFARGGPLNRVLSGRK---IPPDILVNMAVOIARGMNYLLDEAIVPIIHRDLKSSN 263  
QY 109 VVI-----AADGVLKICDFGASRPHNHTHNSLVGTFPMAPEVIGSLPVSETCDTY 160  
Db 264 ILILQKVENGLSNKILKITDFGLAREWHRTTKSAAGTYAMAPAEVIRASMSKSDVW 323  
QY 161 SYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPRSFAELLHQCWEADAKR 220  
Db 324 SYGVLLWELLTGEVPEFRGIDGLAVAYGVAMNKLALPIPTCPEPFAXLMEDCWNPDPHSR 383  
QY 221 PSFKQIISILESM 233  
Db 384 PSFTNILDQLTTI 396

RESULT 10

AAE21717  
ID AAE21717 standard; Protein; 1097 AA.

AC AAE21717;

DT 16-JUL-2002 (first entry)

DE Human PKIN-12 protein.

KW Human; kinase; enzyme; PKIN-12 protein; immune system disorder; anaemia;  
KW acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;  
KW asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;  
KW AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;  
KW leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;  
KW Down's syndrome; gene therapy; protein therapy; cytostatic.

OS Homo sapiens.

XX FH Key Location/Qualifiers  
FT Peptide 1..17  
/label= Signal\_peptide  
FT Protein 18..1097  
/note= "Mature human PKIN-12 protein"

FT Domain 55..114  
/note= "SH3 domain"  
FT Domain 144..403  
/note= "Eukaryotic protein kinase domain"  
FT Domain 146..396  
/note= "Protein kinase domain"  
FT Domain 163..396  
/note= "Protein kinase domain"  
FT Domain 220..233  
/note= "Tyrosine kinase catalytic domain"  
FT Domain 258..276  
/note= "Tyrosine kinase catalytic domain"  
FT Domain 311..321  
/note= "Tyrosine kinase catalytic domain"  
FT Domain 330..352  
/note= "Tyrosine kinase catalytic domain"  
FT Domain 374..396  
/note= "Tyrosine kinase catalytic domain"  
FT Domain 438..749  
/note= "Tyrosine kinase catalytic domain"  
FT Domain 869..893  
/note= "Leucine zipper domain"



DE Human leucine zipper protein kinase.  
XX  
KW Leucine zipper protein kinase; zpk; human; metastasis; tumour;  
KW serine/threonine protein kinase; non-receptor type kinase;  
KW cell hyperproliferation.  
XX  
OS Homo sapiens.  
FH  
FH Key Location/Qualifiers  
FT Protein 1..859  
FT /label= zpk  
FT /note= "leucine zipper protein kinase"  
FT Domain 231..243  
FT /label= protein\_kinase\_domain  
FT /note= "As stated in specification"  
FT Misc-difference 234..235  
FT /note= "Mentioned in specification"  
FT Misc-difference 236..237  
FT /note= "Mentioned in specification"  
FT Misc-difference 240  
FT /note= "Mentioned in specification"  
FT Misc-difference 251  
FT /note= "Mentioned in specification"  
FT Misc-difference 254..256  
FT /note= "Mentioned in specification"  
FT Misc-difference 278..280  
FT /note= "Mentioned in specification"  
FT Misc-difference 292  
FT /note= "Mentioned in specification"  
FT Misc-difference 294..295  
FT /note= "Mentioned in specification"  
FT Misc-difference 297  
FT /note= "Mentioned in specification"  
FT Region 415..418  
FT /note= "Putative endoplasmic reticulum targeting  
FT sequence as given in the specification"  
FT Region 442..468  
FT /label= Leucine\_zipper\_motif  
FT /note= "As stated in specification"  
FT Misc-difference 443  
FT /note= "Mentioned in specification"  
FT Misc-difference 450  
FT /note= "Mentioned in specification"  
FT Misc-difference 457  
FT /note= "Mentioned in specification"  
FT Misc-difference 464  
FT /note= "Mentioned in specification"  
FT /label= ATP\_binding\_site  
FT 537..544  
FT /note= "As stated in specification"  
XX  
PN US5676945-A.  
XX  
PD 14-OCT-1997.  
XX  
PF 01-MAR-1994; 94US-0205018.  
XX  
PR 28-FEB-1995; 95US-0395580.  
PR 01-MAR-1994; 94US-0205018.  
XX  
PA (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.  
XX  
PI Pleasure D, Reddy U;  
XX  
DR WPI; 1997-511822/47.  
DR N-PSDB; AAT89349.  
XX  
PT Human leucine zipper protein kinase - useful for treating tumours of  
PT the central nervous system  
XX  
PS Claim 3; Fig 1; 19pp; English.  
XX  
CC This sequence represents a novel human leucine zipper protein kinase

CC isolated from brain tissue. The specification states that the  
CC protein contains an ATP-binding site at position 537-544 (consensus  
CC sequence Gly-Xaa-Gly-Xaa-Gly), a protein kinase domain at position  
CC 231-243 and a putative endoplasmic reticulum (ER) targeting sequence at  
CC position 413-418 (consensus sequence REEL). This protein is most similar  
CC to members of serine/threonine protein kinases and is believed to be a  
CC "non-receptor type kinase" based on its lack of a transmembrane domain.  
CC Probes to this protein could be used for diagnostic or research purposes  
CC to detect or quantitate the expression of leucine zipper protein kinase.  
CC Overexpression of leucine zipper protein kinase can result in  
CC hyperproliferation of cells and metastasis. The application of exogenous  
CC leucine zipper protein kinase may interfere with specific protein-protein  
CC or protein-nucleic acid interactions involved in hyperproliferation. This  
CC may be used to treat animals suffering from tumours of the central  
CC nervous system by inhibiting the overexpression of leucine zipper protein  
CC kinase in vivo or by interfering with a vital signal in the chain of  
CC signals leading to tumourigenicity.  
XX  
SQ Sequence 859 AA;  
Query Match 37.7%; Score 490; DB 18; Length 859;  
Best Local Similarity 44.3%; Pred. No. 4.9e-42;  
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;  
QY 10 KEAVAKKLKIEKAEI--LSVLSHRNITQFYGVILEPPNYGIVTEVASLGLYDINSN 67  
Db 147 EEVAVKKVRDL-KETDIKHLRLKXHPNITTFKGVCTQAPCYCIIMEFCAQGLYEVLRAG 205  
QY 68 RSEEMDMDHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGVLIKICDFGASR- 126  
Db 206 RPTVPSL-LVDSMGIAAGMNYLHL--KIHRDLKSPNMLTYDDVVKISDFGTSKE 260  
QY 127 FHNHTHMSLVGFPPWMAPEVIOQLPVSETCDTYSYGVVLWEMLTREVPFKGLEQLQAW 186  
Db 261 LSDKSTKMSFAGTVAWMAPEVIRNEPVSERKVDIWSFGVLWELLTGEIPYKDVDSALIW 320  
QY 187 LVVEKNERLTTPSSCPSPFAELHQCEADAKRPSFKQIISLESMSNDT-SLPDK 242  
Db 321 GVGNSLHLPVPSSCPDGFKILLRQCWNSKPRNRPSPRQILHLHDIASADVLSTPQE 377  
RESULT 13  
ABB57049  
ID ABB57049 standard; Protein; 888 AA.  
XX  
AC ABB57049;  
XX  
DT 07-MAR-2002 (first entry)  
XX  
DE Mouse ischaemic condition related protein sequence SEQ ID NO:79.  
XX  
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;  
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.  
XX  
OS Mus musculus.  
XX  
PN WO200188188-A2.  
XX  
PD 22-NOV-2001.  
XX  
PF 18-MAY-2001; 2001WO-JP04192.  
XX  
PR 18-MAY-2000; 2000JP-0145977.  
XX  
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.  
XX  
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;  
XX  
DR WPI; 2002-034733/04.  
DR N-PSDB; ABI99250.  
XX  
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
PT expression levels of particular genes defined in the specification or



PT by determining the expression profile of a gene group comprising these  
PT genes -  
XX  
PS Claim 2; Page 244-248; 2690pp; English.  
XX  
CC The present invention describes a method for examining ischaemic  
CC conditions, comprising measuring the expression levels of particular  
CC genes (I) in a test sample or determining the expression profile of a  
CC gene group in the sample comprising genes selected from (I). The method  
CC is useful for examining the ischaemic condition (e.g. compressive  
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
CC expression levels of particular genes (ABI99202 to ABI99912, encoding  
CC the protein sequences in ABB57020 to ABB57374) or by determining the  
CC expression profile of a gene group comprising these genes. The  
CC expression levels or expression profiles produced by these genes are  
CC used as an indicator when screening for ischaemic condition-improving  
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914  
CC represent PCR primers for a mouse ischaemic condition related sequence,  
CC which are used in the exemplification of the present invention.

Sequence 888 AA;

Query Match 37.6%; Score 489; DB 23; Length 888;  
Best Local Similarity 44.3%; Pred. No. 6.6e-42;  
Matches 105; Conservative 42; Mismatches 80; Indels 10; Gaps 6;

QY 10 KEVAVKLLKIEKEAEI--LSVLSHRNIIQFYGVILEPPNYGIVTEYASLSLYYINSN 67  
Db 180 EEVAVKVKRDL-KETDIKRLKHKHNIITFGVCTQAPCYCIIMEFCAQGQLYEYLRAG 238  
QY 68 RSEMDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGLKICDFGASR- 126  
Db 239 RPYTPSL--LVDWSMGIACGMNYLHLH--KIHRDLKSPNMLITYDDVKISDFGTSKE 293  
QY 127 FHNHTHMSLVGTFPMWAPVIGSLPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAV 186  
Db 294 LSDKSTKMSFAGTVAMWAPVIRNEVSEKVDIWSFGVLMELLTGEIPIKVDVSSAIIW 353  
QY 187 LVVEKNERLITPSSCPRSFAELLHQWEADAKKRPSFKQIISILSMSNDT-SLPDK 242  
Db 354 GVGSNLHLVPSSCPDGFKILLRQCWNTRPRNPSFRQILHLDIASADVLSTPQE 410

RESULT 14  
ABP61000  
ID ABP61000 standard; Protein; 1021 AA.

XX ABP61000;

XX 10-SEP-2002 (first entry)

XX Novel human protein. SEQ ID 87.

XX Human; cytostatic; vulnerary; antiarteriosclerotic; antiparkinsonian;  
KW nootropic; neuroprotective; immunosuppressive; haemostatic;  
KW antiinflammatory; cardiant; antiulcer; virucide; antithyroid;  
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;  
KW wound healing disorders; atherosclerosis; Parkinson's disease;  
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;  
KW inflammation; neoplastic disease; nervous system disorder;  
KW cardiovascular disorders; pancreatitis; respiratory disorder;  
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;  
KW developmental abnormality; gastrointestinal ulceration; neuropathy;  
KW haematological disease; metabolic disease; sperm dysfunction;  
KW thyroid disorder; hypothyroidism; brain damage; colitis;  
KW cone photo- transduction deficiency; neurological disease; stroke;  
KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;  
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;  
KW growth abnormality; precocious puberty.

XX Homo sapiens.

OS  
XX  
PN WO200250105-A1.

XX 27-JUN-2002.  
PD  
XX  
XX 17-DEC-2001; 2001WO-US49232.  
PF  
XX  
XX 19-DEC-2000; 2000US-256710P.  
PR 20-DEC-2000; 2000US-257048P.  
PR 09-JAN-2001; 2001US-260482P.  
PR 30-JAN-2001; 2001US-264922P.  
PR 06-FEB-2001; 2001US-266797P.  
PR 19-MAR-2001; 2001US-276988P.  
PR 04-APR-2001; 2001US-281535P.  
PR 08-MAY-2001; 2001US-289622P.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
PA (GLAX ) GLAXO GROUP LTD.  
XX  
PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y;  
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;  
XX  
XX WPI; 2002-508784/54.  
DR N-PSDB; ABQ86165.  
XX  
XX  
XX Secreted proteins and polynucleotides useful as vaccines for preventing  
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,  
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune  
PT disorder -  
XX  
XX  
XX Claim 1(a); Page 307-309; 335pp; English.

XX The invention relates to an isolated polypeptide with signal sequences  
CC which allow it to be secreted extracellularly or membrane associated.  
CC The activity of polypeptides of the invention may be described as,  
CC cytostatic, vulnerary, antiarteriosclerotic, antiparkinsonian, nootropic,  
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,  
CC cardiant, antiulcer, virucide, antithyroid, cerebroprotective, anorectic,  
CC and metabolic. Polypeptides and polynucleotides of the invention are  
CC useful in the treatment, or as a vaccine in the prevention of, cancer,  
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease  
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,  
CC inflammation, neoplastic diseases, nervous system related disorders and  
CC cardiovascular disorders, systemic autoimmune disease, hyper-immunity,  
CC hyperproliferation, gastrointestinal ulceration, neuropathy,  
CC developmental abnormality, metabolic diseases, sperm dysfunction, thyroid  
CC haematological diseases, metabolic diseases, brain damages, colitis, cone photo-  
CC disorders e.g. hypothyroidism, neurological diseases, stroke, angiogenesis,  
CC transduction deficiency, neurological diseases, thyroid gland, heart,  
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,  
CC growth abnormalities, and alleviation of precocious puberty. The  
CC sequences given in records ABP60965-ABP61019 represent novel human  
CC proteins of the invention.

XX Sequence 1021 AA;

QY Query Match 37.5%; Score 487.5; DB 23; Length 1021;  
Best Local Similarity 40.4%; Pred. No. 1.2e-41;  
Matches 107; Conservative 48; Mismatches 81; Indels 29; Gaps 6;

QY 1 YRAKWSQDKEVAVKLLK-----IEKAEILSVLSHRNIIQFYGVILEPPNY 48  
Db 139 YRATW--QGQEVAVKARQDEQDAAAASVRRERARLFAMLRHPNITELRGVCLQPHL 196  
QY 49 GIVTEYASLSLYDYINSRSEEMDMHIM-TWATDVAKGMHYLHMEAPVKVIHRDLKSR 107  
Db 197 CLVEFARGGALNRALARRIP---PHVLVNWAVQIARGMLYLHEEAFVPIHRDLKSS 252  
QY 108 NVVI-----AADGVLKICDFGASRFHNHTHMSIVGTFPMWAPVIGSLPVSETCDT 159  
Db 253 NILLEKIEHDDICNKTLLKIDFGLAREWHRTTKMSAGTVAMWAPVIRKSGSDI 312  
QY 160 YSYGVVLMEMLTREVPFKGLEGLQVAVMLVVEKNERLITPSSCPRSFAELLHQWEADAKK 219

Db 313 WSYGVLLWEILLTGEVPRYRGIDGLAVYGVAVNKLTLPIPTCEPFAKLMKECWQDDPHI 372  
QY 220 RPSFKQIISILEMSND--TSLPDK 242  
Db 373 RPSFALILEQLTAIEGAVMTMPQE 397

RESULT 15  
AAB85513  
ID AAB85513 standard; protein, 719 AA.  
XX  
AC AAB85513;  
XX  
DT 25-SEP-2001 (first entry)  
XX  
DE Human protein kinase SGK067.  
XX  
KW Protein kinase; enzyme; cytostatic; nootropic; neuroprotective; human; antiparkinsonian; virucide; antibacterial; antifungal; antimigraine; analgesic; hypotensive; hypertensive; immunosuppressive; antiallergic; antipneumatic; antirheumatic; antiarthritic; ophthalmological; anorectic; osteopathic; thrombolytic; antiarteriosclerotic; antiasthmatic; vasotropic; antidiabetic; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200155356-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 25-JAN-2001; 2001WO-US02337.  
XX  
PR 25-JAN-2000; 2000US-0178078.  
PR 31-JAN-2000; 2000US-0179364.  
PR 17-FEB-2000; 2000US-0183173.  
PR 17-MAR-2000; 2000US-0190162.  
PR 29-MAR-2000; 2000US-0193404.  
PR 13-NOV-2000; 2000US-0247013.  
XX  
PA (SUGEN-) SUGEN INC.  
XX  
PI Plowman G, Whyte D, Manning G, Sudarsanam S, Martinez R;  
XX  
DR WPI; 2001-476202/51.  
DR N-PSDB; AAH46913.  
XX  
XX Kinase polypeptides useful for treating cancers, Alzheimer's disease, viral infections, diabetes, obesity, organ transplant rejection and rheumatoid arthritis.  
XX  
PS Claim 7; Page 217; 218pp; English.  
XX  
XX The invention provides human protein kinases and protein kinase-like enzymes and polynucleotides encoding the polypeptides. The kinase polypeptides and their modulators are useful for treating a disease or disorder such as cancer, immune-related diseases, cardiovascular disease, brain or neuronal-associated disease and metabolic disorders, including cancers of tissues, cancers of hematopoietic origin, diseases of the central nervous system, diseases of the peripheral nervous system, Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis, viral infections, infections caused by prions, bacteria and fungi, ocular diseases, migraines, pain, sexual dysfunction, mood disorders, attention disorders, cognitive disorders, hypotension, hypertension, psychotic disorders, neurological disorders, dyskinesias, metabolic disorders, and organ transplant rejection. They are also useful for treating rhinitis, autoimmunity, atherosclerosis, psoriasis, osteoarthritis, asthma, chronic inflammatory pelvic disease, chronic inflammatory bowel disease, rheumatoid arthritis, metabolic disorders such as diabetes, obesity, cardiovascular diseases such as reperfusion injury, coronary thrombosis, clotting disorders and atherosclerosis, ocular diseases such as glaucoma, retinopathy and macular degeneration, psychiatric and neurological disorders such as anxiety, schizophrenia,

CC dementia, manic depression, etc. The polynucleotides are useful in gene  
CC therapy techniques to treat the above mentioned disorders. Sequences  
CC AAB85491-85522 represent the human protein kinases of the invention.  
XX  
SQ Sequence 719 AA;

Query Match 37.2%; Score 483; DB 22; Length 719;  
Best Local Similarity 38.4%; Pred. No. 2.1e-41;  
Matches 106; Conservative 49; Mismatches 85; Indels 36; Gaps 5;  
QY 1 YRAKWSQDKKAVKLLK-----TEKAEILSVLSHRNIIQFYGVILEPPNY 48  
Db 139 YRAW--QGQEVAVKARQDPQDAAAASVREARLFAMLRHPNIIHRGVCLQQPHL 196  
QY 49 GIVTEYASLSLYDYI-----NSNRSEMDMDHMTWATDVAKGMHYLHMEAP 96  
Db 197 CLVLEFARGGALNRLAANAAPDPRAPGPRRARRIPPHVLVNWAVQARGMLYLHBEAF 256  
QY 97 VKVIHRDLKSRNVI-----AADGVLKICDFGASRFHNHTHNSLVGTFPMMAPEVI 148  
Db 257 VPIHRDLKSSNILLLEKIEHDDICNKTLYKTFGLAREWHRRTKSTAGTYAMMAPEVI 316  
QY 149 QSLPVSETCDTYSYGVVLMWLTREVPRKLEGLQVAMLVVEKNERLTISSCPRSAEL 208  
Db 317 KSSLFSKGSIDWSYGVLLWELLTGEVPRYRGIDGLAVYGVAVNKLTLPIPTCEPFAKL 376  
QY 209 LHQCWEADAKRPSFKQIISILEMSND--TSLPDK 242  
Db 377 MKECWQDDPHIRPSFALILEQLTAIEGAVMTMPQE 412

Search completed: May 1, 2003, 20:35:36  
Job time : 50.5397 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:38:08 ; Search time 15.8466 Seconds

(without alignments)  
1344.947 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_31\_277  
Perfect score: 1300  
Sequence: 1 YRAKWSQDKEVAVKKLKI.....SILESMSNDTSLPDKNSFL 247

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

a1 number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US05\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID                 | Description        |
|------------|-------|-------------|--------|-----------------------|--------------------|
| 1          | 1300  | 100.0       | 455    | 10 US-09-757-982-5    | Sequence 5, Appl1  |
| 2          | 510.5 | 39.3        | 394    | 10 US-09-862-027-19   | Sequence 19, Appl1 |
| 3          | 495.5 | 38.1        | 328    | 10 US-09-862-027-18   | Sequence 18, Appl1 |
| 4          | 483   | 37.2        | 1036   | 12 US-10-014-882-2    | Sequence 2, Appl1  |
| 5          | 482.5 | 37.1        | 847    | 9 US-10-143-133-2     | Sequence 2, Appl1  |
| 6          | 469.5 | 36.1        | 966    | 10 US-09-771-161A-197 | Sequence 197, App  |
| 7          | 448.5 | 34.5        | 850    | 10 US-09-904-389-2    | Sequence 2, Appl1  |
| 8          | 442   | 34.0        | 263    | 10 US-09-840-704-5    | Sequence 5, Appl1  |
| 9          | 424   | 32.6        | 92     | 9 US-09-764-868-799   | Sequence 799, App  |
| 10         | 395   | 30.4        | 579    | 9 US-10-158-895-4     | Sequence 4, Appl1  |
| 11         | 395   | 30.4        | 590    | 9 US-10-158-895-15    | Sequence 15, Appl1 |
| 12         | 351.5 | 27.0        | 251    | 8 US-08-987-689A-32   | Sequence 32, Appl1 |
| 13         | 343   | 26.4        | 505    | 9 US-09-977-260-6     | Sequence 6, Appl1  |
| 14         | 343   | 26.4        | 505    | 9 US-09-977-261-6     | Sequence 6, Appl1  |
| 15         | 343   | 26.4        | 505    | 10 US-09-977-269-6    | Sequence 6, Appl1  |
| 16         | 343   | 26.4        | 505    | 10 US-09-982-610-20   | Sequence 20, Appl1 |
| 17         | 336   | 25.8        | 1036   | 10 US-09-771-161A-255 | Sequence 255, App  |
| 18         | 336   | 25.8        | 1036   | 10 US-09-771-161A-256 | Sequence 256, App  |
| 19         | 335.5 | 25.8        | 537    | 9 US-09-977-260-11    | Sequence 11, Appl1 |

|    |       |      |     |                       |                    |
|----|-------|------|-----|-----------------------|--------------------|
| 20 | 335.5 | 25.8 | 537 | 9 US-09-977-261-11    | Sequence 11, Appl1 |
| 21 | 335.5 | 25.8 | 537 | 10 US-09-977-269-11   | Sequence 11, Appl1 |
| 22 | 335.5 | 25.8 | 835 | 10 US-09-947-199-8    | Sequence 8, Appl1  |
| 23 | 335   | 25.8 | 764 | 10 US-09-925-302-714  | Sequence 714, App  |
| 24 | 333.5 | 25.7 | 537 | 10 US-09-771-161A-212 | Sequence 212, App  |
| 25 | 333.5 | 25.7 | 537 | 10 US-09-771-161A-213 | Sequence 213, App  |
| 26 | 332   | 25.5 | 310 | 9 US-09-939-833-7     | Sequence 7, Appl1  |
| 27 | 332   | 25.5 | 310 | 10 US-09-939-754-7    | Sequence 7, Appl1  |
| 28 | 332   | 25.5 | 310 | 10 US-09-939-832-7    | Sequence 7, Appl1  |
| 29 | 332   | 25.5 | 425 | 10 US-09-828-313-29   | Sequence 29, Appl1 |
| 30 | 332   | 25.5 | 822 | 9 US-09-757-415A-2    | Sequence 2, Appl1  |
| 31 | 328.5 | 25.3 | 450 | 9 US-09-977-260-7     | Sequence 7, Appl1  |
| 32 | 328.5 | 25.3 | 450 | 9 US-09-977-261-7     | Sequence 7, Appl1  |
| 33 | 328.5 | 25.3 | 450 | 10 US-09-977-269-7    | Sequence 7, Appl1  |
| 34 | 328.5 | 25.3 | 835 | 10 US-09-947-199-2    | Sequence 2, Appl1  |
| 35 | 327.5 | 25.2 | 258 | 10 US-09-840-704-3    | Sequence 3, Appl1  |
| 36 | 324.5 | 25.0 | 536 | 9 US-09-977-260-13    | Sequence 13, Appl1 |
| 37 | 324.5 | 25.0 | 536 | 9 US-09-929-266-10    | Sequence 10, Appl1 |
| 38 | 324.5 | 25.0 | 536 | 9 US-09-977-261-13    | Sequence 13, Appl1 |
| 39 | 324.5 | 25.0 | 536 | 10 US-09-977-269-13   | Sequence 13, Appl1 |
| 40 | 324   | 24.9 | 536 | 9 US-09-977-260-12    | Sequence 12, Appl1 |
| 41 | 324   | 24.9 | 536 | 9 US-09-977-261-12    | Sequence 12, Appl1 |
| 42 | 324   | 24.9 | 536 | 10 US-09-977-269-12   | Sequence 12, Appl1 |
| 43 | 321   | 24.7 | 256 | 10 US-09-840-704-4    | Sequence 4, Appl1  |
| 44 | 320   | 24.6 | 543 | 9 US-09-977-260-14    | Sequence 14, Appl1 |
| 45 | 320   | 24.6 | 543 | 9 US-09-977-261-14    | Sequence 14, Appl1 |

ALIGNMENTS

RESULT 1  
US-09-757-982-5  
Sequence 5, Application US/09757982  
Patent No. US20020094559A1  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: MNI-050  
CURRENT APPLICATION NUMBER: US/09/757,982  
PRIOR FILING DATE: 2001-01-10  
PRIOR APPLICATION NUMBER: 09/163,115  
PRIOR FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-757-982-5

|                       |                       |                         |                  |                             |
|-----------------------|-----------------------|-------------------------|------------------|-----------------------------|
| Query Match           | 100.0%                | Score 1300;             | DB 10;           | Length 455;                 |
| Best Local Similarity | 100.0%                | Pred. No. 1e-98;        |                  |                             |
| Matches 247;          | Conservative 0;       | Mismatches 0;           | Indels 0;        | Gaps 0;                     |
| QY                    | 1 YRAKWSQDKEVAVKKLKI  | EAEILSVLSHRNIIQFYGVILEP | PNYGI            | TEVASLGL 60                 |
| DB                    | 31 YRAKWSQDKEVAVKKLKI | EAEILSVLSHRNIIQFYGVILEP | PNYGI            | TEVASLGL 90                 |
| QY                    | 61 YDYINSNRSEEMDMHIM  | TWATDVAKGMHYLHMEAPVYI   | HRDLKSRNV        | VIADGVLIK 120               |
| DB                    | 91 YDYINSNRSEEMDMHIM  | TWATDVAKGMHYLHMEAPVYI   | HRDLKSRNV        | VIADGVLIK 150               |
| QY                    | 121 DFGASRFHNHTTHMSLV | GFPPMAPEVIO             | SLPVSE           | CTDYSYGVLWEMLTREVPFKGLE 180 |
| DB                    | 151 DFGASRFHNHTTHMSLV | GFPPMAPEVIO             | SLPVSE           | CTDYSYGVLWEMLTREVPFKGLE 210 |
| QY                    | 181 GLOYAWLVEKNERLT   | IPSSCP                  | PSFAELLHQCWEADAK | KRPSPKQIISILESMSNDTSLP 240  |
| DB                    | 211 GLOYAWLVEKNERLT   | IPSSCP                  | PSFAELLHQCWEADAK | KRPSPKQIISILESMSNDTSLP 270  |
| QY                    | 241 DKCNSFL 247       |                         |                  |                             |





```
; Sequence 2, Application US/10143133
; Publication No. US20020197658A1
; GENERAL INFORMATION:
; APPLICANT: Yoganathan, Thillainathan
; APPLICANT: Delaney, Allen
; TITLE OF INVENTION: Cancer Associated Protein Kinase and Its Use
; FILE REFERENCE: KINE-023
; CURRENT APPLICATION NUMBER: US/10/143,133
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 60/290,555
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-143-133-2
```

```
Query Match 37.1%; Score 482.5; DB 9; Length 847;
Best Local Similarity 41.1%; Pred. No. 1.5e-31;
Matches 104; Conservative 44; Mismatches 80; Indels 25; Gaps 6;
```

```
QY 1 YRAKWISQ-----DKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYG 49
Db 132 YRGSWRGELVAVKARQDPDEDISV-TAESVROEARLFAMLAHPNIIALKAVCLEEPNLC 190
QY 50 IVTEYASIGSLYDINSNRSEEMDMHIM-TWATDVAKGMHYLHMEAPVKVIHRDLKSRN 108
Db 191 LVMEYVAGGFLSRALAGRVP-----PHVLVNMVAVQIARGMYLHCEALVPIHRDLKSN 246
QY 109 VV-----IAADGV---LKICDFGASRFHNHTTHMSLVGTFPWMAPEVIQSLPVSETCTY 160
Db 247 ILLQPIESDDMEHKTITKIDFGLAEMHKTQMSAAGTYAMMAPEVIKASTFSKSDVW 306
QY 161 SYGVVLEMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELLHQCEWADAKKR 220
Db 307 SFGVLMELLTGEVPRGIDCLAVAVGVAVNKLTLPIPTCTPEBPFAQLMADCWADPHRR 366
QY 221 PSFKQIISILESM 233
Db 367 PDFASILQOLEAL 379
```

```
RESULT 6
US-09-771-161A-197
; Sequence 197, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 197
; LENGTH: 966
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-771-161A-197
```

```
Query Match 36.1%; Score 469.5; DB 10; Length 966;
Best Local Similarity 42.2%; Pred. No. 2e-30;
Matches 97; Conservative 47; Mismatches 77; Indels 9; Gaps 5;
```

```
QY 10 KEVAVKLLKIEKAEI--LSVLSHRNIIQFYGVILEPPNYGIVTEYASIGSLYDINSN 67
```

```
Db 190 EEVAIKK-VREQNETDIKHLRKLKHPNIIAFKGVCTQAPCYCIIMEYCAHGQLEYVLRA 248
QY 68 RSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVIIAADVLKICDFGASR- 126
Db 249 R--KITPRLVDWSTGASGMNYLHLH--KIIHRDLKSPNVLTHTDAVKISDFGTSKE 303
QY 127 FHNHTTHMSLVGTFPWMAPEVIQSLPVSETCTPYSGVVLWMLTREVPFKGLEGLQVAM 186
Db 304 LSDKSTKMSFAGTVAMMAPEVIRNEPVSSEKVDIWSFGVLMELLTGEIPYKDVDSALIW 363
QY 187 LVEKNERLTIPSSCPSPFAELLHQCEWADAKKRPSFKQIISILESMND 236
Db 364 GVGNSLSLHPVPSTCPDGFKILMKQTWSKPRNRPFRQTLMLHLDIASAD 413
```

```
RESULT 7
US-09-904-389-2
; Sequence 2, Application US/09904389
; Patent No. US20020129404A1
; GENERAL INFORMATION:
; APPLICANT: Clendennen, Stephanie K.
; APPLICANT: Schuster, Debra K.
; TITLE OF INVENTION: CTR1 HOMOLOGUE FROM MELON
; FILE REFERENCE: 4257-0029.30
; CURRENT APPLICATION NUMBER: US/09/904,389
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/218,307
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Cucumis melo
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (154)...(154)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-09-904-389-2
```

```
Query Match 34.5%; Score 448.5; DB 10; Length 850;
Best Local Similarity 40.2%; Pred. No. 9e-29;
Matches 99; Conservative 38; Mismatches 94; Indels 15; Gaps 4;
```

```
QY 1 YRAKWISQDKEVAVKLLKIE-----KEAEILSVLSHRNIIQFYGVILEPPNYGI 50
Db 592 YRGEMHGS--VAVKILTEQDFHERVNEFLREVAIMKSLHPNIVLFMGAVTKPPNLSI 649
QY 51 VTEYASIGSLYDINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNV 110
Db 650 VTEYLSRGSLLYRLHKSGVKDIDETRINMAFDVAKGMNYLHRDP-PIVHRDLKSPNLL 708
QY 111 IAADGVKICDFGASRFHNHT--THMSLVGTFPWMAPEVIQSLPVSETCTYSGVVLWE 168
Db 709 VDKKYYTVKVCDFGLSRKARTFLSSKSAAGTPEWMAPEVLARDEPSNEKSDVYSFGVILME 768
QY 169 MLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELLHQCEWADAKKRPSFKQIIS 228
Db 769 LATIQOPWCNLPQAQVAAVGFGRKRLDIPRDVNPKLASLIIVACWADEPWKRPSFSIME 828
QY 229 ILESMS 234
Db 829 TLKPMT 834
```

```
RESULT 8
US-09-840-704-5
; Sequence 5, Application US/09840704
; Patent No. US20020122801A1
; GENERAL INFORMATION:
; APPLICANT: Dedhar, Shoukat
; APPLICANT: Hannigan, Greg
```

TITLE OF INVENTION: Intergrin-Linked Kinase and its Uses  
FILE REFERENCE: KIN-2CON  
CURRENT APPLICATION NUMBER: US/09/840,704  
CURRENT FILING DATE: 2001-04-23  
PRIOR APPLICATION NUMBER: 09/566,906  
PRIOR FILING DATE: 2000-05-09  
PRIOR APPLICATION NUMBER: US08/752,345  
PRIOR FILING DATE: 1996-11-19  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 263  
TYPE: PRT  
ORGANISM: H. sapiens  
FEATURE:  
NAME/KEY: Other  
LOCATION: (1)...(263)  
US-09-840-704-5

Query Match 34.0%; Score 442; DB 10; Length 263;  
Best Local Similarity 40.2%; Pred. No. 8.1e-29;  
Matches 99; Conservative 41; Mismatches 90; Indels 16; Gaps 5;

QY 1 YRAKVISQDKVAVKLLKIE-----KEAELLSVLSHRNIIQFYGVILEPPNYGI 50  
DB 21 HRAEWHGSD--VAVKILMEQDFHAEVNEFLREVALIMKRLHPNIVLFMGAVTQPNLSI 78  
QY 51 VTEYASLSGLYDYN-SNRSEEMDMHMTWATDVAKGMHYLMEAPVKVIHRDLKSRNV 109  
DB 79 VTEYLSRGSLSYRLHLKSGAREQLDERRLSMAYDVAKGMNLIHNRNP-PIVHRDLKSPNL 137  
QY 110 VIAADGVLCIDFGASRFHNHT--THMSLVGTFPMAPEVIOQLPVSETCDTYSYGVLM 167  
DB 138 LVDKKYTVKVCDFGLSRKASTPLSSKSAAGTPEWMAPEVLDEPSNEKSDVYSGVILM 197  
QY 168 EMLTREVPFKGLEGLQVAVLVEKNERLTISSCPSPFAELLHQWEADAKRPSFKQII 227  
DB 198 ELATLQPGWNLNPAQVAAVGVFKCKRLEIPRLNPQVAAIEGCTNEPWRKPSFATIM 257  
QY 228 SILES 233  
DB 258 DLRLPL 263

RESULT 9  
US-09-764-868-799  
Sequence 799, Application US/09764868  
Patent No. US20020168711A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT232  
CURRENT APPLICATION NUMBER: US/09/764,868  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 1510  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 799  
LENGTH: 92  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (4)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-868-799

Query Match 32.6%; Score 424; DB 9; Length 92;  
Best Local Similarity 98.8%; Pred. NO. 7.3e-28;  
Matches 80; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 23 EAEILSVLSHRNIIQFYGVILEPPNYGI VTEYASLSGLYDYN-SNRSEEMDMHMTWAT 82  
:|||||

DB 12 QAEILSVLSHRNIIQFYGVILEPPNYGI VTEYASLSGLYDYN-SNRSEEMDMHMTWAT 71  
QY 83 DVAKGMHYLMEAPVKVIHRD 103  
DB 72 DVAKGMHYLMEAPVKVIHRD 92

RESULT 10  
US-10-158-895-4  
Sequence 4, Application US/10158895  
Patent No. US20020155624A1  
GENERAL INFORMATION:  
APPLICANT: ONO, KOICHIRO  
APPLICANT: OHTOMO, TOSHIHIKO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/10/158,895  
CURRENT FILING DATE: 2002-06-03  
PRIOR APPLICATION NUMBER: US/09/529,279  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188  
PRIOR FILING DATE: 1997-10-22  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 579  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-158-895-4

Query Match 30.4%; Score 395; DB 9; Length 579;  
Best Local Similarity 37.1%; Pred. No. 1.4e-24;  
Matches 91; Conservative 46; Mismatches 80; Indels 28; Gaps 11;

QY 2 RAKVISQDKVAVKLLKIEKAE-----LSVLSHRNIIQFYGVILEPPNYGI 52  
DB 52 KAKM--RAKDVAIK---QIESESERKAFIVELRLQLSRVNHPNIVKLYGACLP--VCLVM 104  
QY 53 EYASLSGLYDYN-SNRSEEM--DMHMTWATDVAKGMHYLMEAPVKVIHRDLKSRNV 109  
DB 105 EYAEGLSYLVNVLHG--AEPLPYTAAHAMSWCLQCSQGVAYLHSMQKALIHRLKPPNL 162  
QY 110 VIAADG-VLKICDFG-ASRFHNHTTHMSLVGTFPMAPEVIOQLPVSETCDTYSYGVLM 167  
DB 163 LVVAGTVLKICDFGTACDIQTHMTNNK--GSAAMWMAPEVLEGSNYSERKCDVFSWGIILM 220  
QY 168 EMLTREVPFKGLEG--LOVAVLVEKNERLTISSCPSPFAELLHQWEADAKRPSFKQ 225  
DB 221 EVITRRKPFDEIGPARIRMW-AVHNGTRPPLKLPKPIESLMTKCSKDPSPSME 279  
QY 226 IISIL 230  
DB 280 IVKIM 284

RESULT 11  
US-10-158-895-15  
Sequence 15, Application US/10158895  
Patent No. US20020155624A1  
GENERAL INFORMATION:  
APPLICANT: ONO, KOICHIRO  
APPLICANT: OHTOMO, TOSHIHIKO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/10/158,895  
CURRENT FILING DATE: 2002-06-03  
PRIOR APPLICATION NUMBER: US/09/529,279  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796

PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188  
PRIOR FILING DATE: 1997-10-22  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 15  
LENGTH: 590  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-158-895-15

Query Match 30.4%; Score 395; DB 9; Length 590;  
Best Local Similarity 37.1%; Pred. No. 1.4e-24;  
Matches 91; Conservative 46; Mismatches 80; Indels 28; Gaps 11;

QY 2 RAKWISQDKEAVAKLKEAE-----ILSVLSHRNIIQFYGVILEPPNYGIYT 52  
DB 52 KAKM--RAKDVAIK---QIESESERKAFIVELRQLSRVNHPIVKLYGACLP--VCLVM 104  
53 EYASLSLYDYINSNRSEEM---DMDHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNV 109  
105 EYAEGLSLYVNLHG--AEPLPYTAAHANSMWCLQCSQGVAYLHSMQPKALIHRLKPPNL 162  
QY 110 VIAADG-VLKICDFG-ASRFHNHTHMSLVGTFPMAPEVIOQLPVSETCDTYSYGVVLM 167  
DB 163 LLVAGGTVLKICDFGTACDIQTHMTNNK--GSAAMWAPFVFGSNYSKCDVFSWGILW 220  
QY 168 EMLTREVPEFKLEG--LQVAVLVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSFKQ 225  
DB 221 EVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPSMEE 279  
QY 226 IISIL 230  
DB 280 IVKIM 284

## RESULT 12

US-08-987-689A-32  
Sequence 32, Application US/08987689A  
Patent No. US20020048782A1  
GENERAL INFORMATION:

APPLICANT: Sima Lev  
APPLICANT: Joseph Schlesinger  
TITLE OF INVENTION: PYK2 RELATED PRODUCTS AND METHODS  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA

ZIP: 90071

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/987,689A

FILING DATE: December 9, 1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/032,824

FILING DATE: December 11, 1996

APPLICATION NUMBER: 08/460,626

FILING DATE: June 2, 1995

APPLICATION NUMBER: 08/357,642

FILING DATE: December 15, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 230/110

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-987-689A-32

Query Match 27.0%; Score 351.5; DB 8; Length 251;  
Best Local Similarity 32.9%; Pred. No. 1.9e-21;  
Matches 78; Conservative 50; Mismatches 94; Indels 15; Gaps 7;

QY 1 YRAKWISQDKEAVAKL-----LKE---KEAELLSVLSHRNIIQFYGVILEPPNYGIYT 53  
DB 15 YEGWKKYSLTVAVKTLKEDTMEVEEFLKEAVMKEIKHPNLVQLGVCRTREPPFIITE 74  
QY 54 YASLSLYDYINSNRSEEMDMDHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVIAA 113  
DB 75 FMTYGNLDYLRRCNRQEVNAVVLVMTQISSAMEYLEK--NFIHRDLAARNCLVGE 131  
QY 114 DGLKICDFGASRFHNHTHMSLVGT-FP--WMAPEVIOQLPVSETCDTYSYGVVLEML 170  
DB 132 NHLKVADFGLSRLMTGDTYTAHAGAKPIKWTAPESLAYNKFISDVWAFGLWLEIA 191  
QY 171 TREV-PFKGLEGLQVAVLVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSFKQI 226  
DB 192 TYGMSPPFGIDRSQV-YELLEKDYMKRPEGCEPKVETLMRACWQWNPSPDRPSFAI 247

## RESULT 13

US-09-977-260-6

Sequence 6, Application US/09977260

Publication No. US20020192790A1

GENERAL INFORMATION:

APPLICANT: ULLRICH, AXEL

APPLICANT: GISHIZKY, MIKHAIL

APPLICANT: SURES, IRMINGARD

TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES

FILE REFERENCE: 038602/1260

CURRENT APPLICATION NUMBER: US/09/977,260

PRIOR FILING DATE: 2001-10-16

PRIOR APPLICATION NUMBER: 08/232,545

PRIOR FILING DATE: 1994-04-22

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 6

LENGTH: 505

TYPE: PRT

ORGANISM: Unknown Organism

FEATURE:

OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte

US-09-977-260-6

Query Match 26.4%; Score 343; DB 9; Length 505;  
Best Local Similarity 33.7%; Pred. No. 2.1e-20;  
Matches 85; Conservative 50; Mismatches 95; Indels 22; Gaps 9;

QY 12 VAVKKL-----KIEKEAELLSVLSHRNIIQFYGV-ILEPPNYGIYVEYASLSGLYDY 63  
DB 259 VAVKTLKPGSMDPNDFLREAQIMKNLRHPKLIQLYAVCTLEDPIY-ITELMRHGSLOEY 317  
QY 64 INSNRSEEMDMDHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVIAADGLKICDFG 123  
DB 318 LQNDTGSKIHLTQYDMAQAASGMAVLESR--NYIHRDLAARNVLVGEHNIYKVADFG 374  
QY 124 ASRFHN-----HTHMSLVGTFPMAPEVIOQLPVSETCDTYSYGVVLEMLT-REVPF 176  
DB 375 LARVFKVDNEDIYESRHEIKLPVKWTAPEAIRSNKFSIKSDVMSFGILLYEITTYGMPY 434



OY 177 KGLEGLQVAMLVVEKNERLTIPSSCPSPFAELLHQCEWADAKRPSFKQIISILES-MSN 235  
DB 435 SGMGTGAQVIQMLAQ-NYRLPQPSNCPQGFYNNIMLECWNAEPKERPTFETLRWKLEDYFET 493  
OY 236 DTSLPDKCNSFL 247  
DB 494 DSSYSD-ANNFI 504

RESULT 14  
US-09-977-261-6  
; Sequence 6, Application US/09977261  
; Publication No. US20030054527A1

; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, AXEL  
; APPLICANT: GISHIZKY, MIKHAEL  
; APPLICANT: SURES, IRMINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1259  
; CURRENT APPLICATION NUMBER: US/09/977,261  
; PRIOR FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 08/232,545  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 505  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte  
; OTHER INFORMATION: kinase 3  
US-09-977-261-6

Query Match 26.4%; Score 343; DB 9; Length 505;  
Best Local Similarity 33.7%; Pred. No. 2.1e-20;  
Matches 85; Conservative 50; Mismatches 95; Indels 22; Gaps 9;

OY 12 VAVKKLL-----KIEKAEILSVLSHRNIIQFYGV-ILEPPNYGIVTEYASLSGLYDY 63  
DB 259 VAVKTLKPGSMDPNDPLREAQIMKNLRHKLQLYAVCTLEDPIY-ITTELMRHGSLQEY 317  
OY 64 INSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIADGVLCIDFG 123  
DB 318 LQNDTGSKIHLLTQVDMAQVAGMAYLESR---NYIHRDLAARNVLVGEHNIYKVADFG 374  
124 ASRFHN-----HTTHMSLVGTFPMAPEVIQSLPVSETCDTYSYGVVLMEMLT-REVPF 176  
375 LARVFKVDNEDIYESRHEIKLPVKWTAPEAIRSNKFSIKSDVMSFGILLYEITTYGKMPY 434  
OY 177 KGLEGLQVAMLVVEKNERLTIPSSCPSPFAELLHQCEWADAKRPSFKQIISILES-MSN 235  
DB 435 SGMGTGAQVIQMLAQ-NYRLPQPSNCPQGFYNNIMLECWNAEPKERPTFETLRWKLEDYFET 493  
OY 236 DTSLPDKCNSFL 247  
DB 494 DSSYSD-ANNFI 504

RESULT 15  
US-09-977-269-6  
; Sequence 6, Application US/09977269  
; Patent No. US20020082037A1

; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, AXEL  
; APPLICANT: GISHIZKY, MIKHAEL  
; APPLICANT: SURES, IRMINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1260  
; CURRENT APPLICATION NUMBER: US/09/977,269  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 08/232,545

; PRIOR FILING DATE: 1994-04-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 505  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte  
US-09-977-269-6

Query Match 26.4%; Score 343; DB 10; Length 505;  
Best Local Similarity 33.7%; Pred. No. 2.1e-20;  
Matches 85; Conservative 50; Mismatches 95; Indels 22; Gaps 9;

OY 12 VAVKKLL-----KIEKAEILSVLSHRNIIQFYGV-ILEPPNYGIVTEYASLSGLYDY 63  
DB 259 VAVKTLKPGSMDPNDPLREAQIMKNLRHKLQLYAVCTLEDPIY-ITTELMRHGSLQEY 317  
OY 64 INSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIADGVLCIDFG 123  
DB 318 LQNDTGSKIHLLTQVDMAQVAGMAYLESR---NYIHRDLAARNVLVGEHNIYKVADFG 374  
OY 124 ASRFHN-----HTTHMSLVGTFPMAPEVIQSLPVSETCDTYSYGVVLMEMLT-REVPF 176  
DB 375 LARVFKVDNEDIYESRHEIKLPVKWTAPEAIRSNKFSIKSDVMSFGILLYEITTYGKMPY 434  
OY 177 KGLEGLQVAMLVVEKNERLTIPSSCPSPFAELLHQCEWADAKRPSFKQIISILES-MSN 235  
DB 435 SGMGTGAQVIQMLAQ-NYRLPQPSNCPQGFYNNIMLECWNAEPKERPTFETLRWKLEDYFET 493  
OY 236 DTSLPDKCNSFL 247  
DB 494 DSSYSD-ANNFI 504

Search completed: May 1, 2003, 20:52:14  
Job time : 17.8466 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:33:12 ; Search time 16.5356 Seconds  
(without alignments)  
439.505 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_31\_277  
Perfect score: 1300  
Sequence: 1 YRAKWSQDKVAVKLLKLT.....SILESMSNDTSLPDKNSFL 247

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID                  | Description        |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1          | 1300  | 100.0       | 455    | 3 US-09-221-235-5   | Sequence 5, Appl1  |
| 2          | 1300  | 100.0       | 455    | 3 US-09-221-928-5   | Sequence 5, Appl1  |
| 3          | 1300  | 100.0       | 455    | 4 US-09-221-527-5   | Sequence 5, Appl1  |
| 4          | 1300  | 100.0       | 455    | 4 US-09-221-236-5   | Sequence 5, Appl1  |
| 5          | 1300  | 100.0       | 455    | 4 US-09-221-416-5   | Sequence 5, Appl1  |
| 6          | 1300  | 100.0       | 455    | 4 US-09-221-245-5   | Sequence 5, Appl1  |
| 7          | 1300  | 100.0       | 455    | 4 US-09-163-115-5   | Sequence 5, Appl1  |
| 8          | 1300  | 100.0       | 455    | 4 US-09-221-528-5   | Sequence 5, Appl1  |
| 9          | 1300  | 100.0       | 455    | 4 US-09-593-553-5   | Sequence 5, Appl1  |
| 10         | 1300  | 100.0       | 455    | 4 US-09-221-237-5   | Sequence 5, Appl1  |
| 11         | 490   | 37.7        | 668    | 1 US-08-205-018-2   | Sequence 2, Appl1  |
| 12         | 490   | 37.7        | 859    | 1 US-08-395-580-2   | Sequence 2, Appl1  |
| 13         | 490   | 37.7        | 859    | 5 PCT-US95-02792-2  | Sequence 2, Appl1  |
| 14         | 442   | 34.0        | 263    | 3 US-09-035-706-5   | Sequence 5, Appl1  |
| 15         | 442   | 34.0        | 263    | 3 US-08-955-841-5   | Sequence 5, Appl1  |
| 16         | 442   | 34.0        | 263    | 4 US-09-390-425-5   | Sequence 5, Appl1  |
| 17         | 442   | 34.0        | 263    | 4 US-09-566-906-5   | Sequence 5, Appl1  |
| 18         | 442   | 34.0        | 821    | 1 US-07-928-464-2   | Sequence 2, Appl1  |
| 19         | 442   | 34.0        | 821    | 1 US-08-003-311B-2  | Sequence 2, Appl1  |
| 20         | 442   | 34.0        | 821    | 1 US-08-261-432-2   | Sequence 2, Appl1  |
| 21         | 442   | 34.0        | 821    | 5 PCT-US93-07347-2  | Sequence 2, Appl1  |
| 22         | 395   | 30.4        | 579    | 4 US-09-529-279-4   | Sequence 4, Appl1  |
| 23         | 395   | 30.4        | 590    | 4 US-09-529-279-15  | Sequence 15, Appl1 |
| 24         | 363.5 | 28.0        | 275    | 2 US-08-701-191A-36 | Sequence 36, Appl1 |
| 25         | 359.5 | 27.7        | 261    | 2 US-07-857-224B-59 | Sequence 59, Appl1 |
| 26         | 356.5 | 27.4        | 261    | 2 US-07-857-224B-60 | Sequence 60, Appl1 |
| 27         | 355.5 | 27.3        | 1584   | 4 US-09-457-040B-27 | Sequence 27, Appl1 |

|    |       |      |     |                     |                    |
|----|-------|------|-----|---------------------|--------------------|
| 28 | 343   | 26.4 | 505 | 1 US-08-222-616-20  | Sequence 20, Appl1 |
| 29 | 343   | 26.4 | 505 | 4 US-08-446-648-20  | Sequence 20, Appl1 |
| 30 | 343   | 26.4 | 505 | 5 PCT-US95-04228-20 | Sequence 20, Appl1 |
| 31 | 341   | 26.2 | 506 | 4 US-08-426-509A-6  | Sequence 6, Appl1  |
| 32 | 341   | 26.2 | 511 | 5 PCT-US95-05008-6  | Sequence 6, Appl1  |
| 33 | 339   | 26.1 | 820 | 1 US-08-166-717D-6  | Sequence 6, Appl1  |
| 34 | 335.5 | 25.8 | 259 | 2 US-07-857-224B-52 | Sequence 52, Appl1 |
| 35 | 335.5 | 25.8 | 537 | 4 US-08-426-509A-11 | Sequence 11, Appl1 |
| 36 | 335.5 | 25.8 | 537 | 5 PCT-US95-05008-11 | Sequence 11, Appl1 |
| 37 | 335   | 25.8 | 313 | 1 US-08-278-089A-17 | Sequence 17, Appl1 |
| 38 | 335   | 25.8 | 313 | 2 US-08-838-957A-16 | Sequence 16, Appl1 |
| 39 | 335   | 25.8 | 729 | 1 US-07-640-029-3   | Sequence 3, Appl1  |
| 40 | 335   | 25.8 | 731 | 1 US-07-921-807B-5  | Sequence 5, Appl1  |
| 41 | 335   | 25.8 | 731 | 1 US-08-441-944A-5  | Sequence 3, Appl1  |
| 42 | 335   | 25.8 | 731 | 4 US-08-439-992A-3  | Sequence 3, Appl1  |
| 43 | 335   | 25.8 | 733 | 1 US-07-640-029-4   | Sequence 4, Appl1  |
| 44 | 335   | 25.8 | 733 | 1 US-07-921-807B-6  | Sequence 6, Appl1  |
| 45 | 335   | 25.8 | 733 | 1 US-08-441-944A-6  | Sequence 6, Appl1  |

ALIGNMENTS

RESULT 1  
US-09-221-235-5  
; Sequence 5, Application US/09221235  
; Patent No. 6043040  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: NMI-050  
; CURRENT APPLICATION NUMBER: US/09/221,235  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-221-235-5

Query Match 100.0%; Score 1300; DB 3; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |   |     |
|----|-----|---|-----|
| QY | 1   | YRAKWSQDKVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPBNYGIIVTEYASLSGL  | 60  |
| DB | 31  | YRAKWSQDKVAVKLLKIEKEAELISVLSHRNIIQFYGVILEPBNYGIIVTEYASLSGL  | 90  |
| QY | 61  | YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVIIAADGVLKIC | 120 |
| DB | 91  | YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVIIAADGVLKIC | 150 |
| QY | 121 | DFGASRFNHTTMSLVGTFPMWABEVIQSLPVSETCDTYSYGVVLMMLTREVPFKGLE   | 180 |
| DB | 151 | DFGASRFNHTTMSLVGTFPMWABEVIQSLPVSETCDTYSYGVVLMMLTREVPFKGLE   | 210 |
| QY | 181 | GLQVAMLVKEKNERLTIPSSCPSPFAELLHQCEADAKKRPSPKQIISILESMSNDTSLP | 240 |
| DB | 211 | GLQVAMLVKEKNERLTIPSSCPSPFAELLHQCEADAKKRPSPKQIISILESMSNDTSLP | 270 |
| QY | 241 | DKCNSTFL 247  |     |
| DB | 271 | DKCNSTFL 277  |     |

RESULT 2  
US-09-221-928-5  
; Sequence 5, Application US/09221928  
; Patent No. 6121030  
; GENERAL INFORMATION:

APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,928  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-928-5

Query Match 100.0%; Score 1300; DB 3; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRAKWSQDKEVAVKKLKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60  
DB 31 YRAKWSQDKEVAVKKLKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90  
OY 61 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120  
DB 91 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150  
OY 121 DFGASRFHNHTTHMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLE 180  
DB 151 DFGASRFHNHTTHMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLE 210  
OY 181 GLQVAMLVVEKNERLTIPSSCPSPFAELLHQCWEADAKKRPSEFKQIISILESMSNDTSLP 240  
DB 211 GLQVAMLVVEKNERLTIPSSCPSPFAELLHQCWEADAKKRPSEFKQIISILESMSNDTSLP 270  
OY 241 DKCNSFL 247  
DB 271 DKCNSFL 277

RESULT 3  
US-09-221-527-5  
; Sequence 5; Application US/09221527  
; Patent No. 6146832  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,527  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-527-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRAKWSQDKEVAVKKLKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60  
DB 31 YRAKWSQDKEVAVKKLKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90  
OY 61 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120  
DB 91 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150  
OY 121 DFGASRFHNHTTHMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLE 180

DB 151 DFGASRFHNHTTHMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLE 210  
OY 181 GLQVAMLVVEKNERLTIPSSCPSPFAELLHQCWEADAKKRPSEFKQIISILESMSNDTSLP 240  
DB 211 GLQVAMLVVEKNERLTIPSSCPSPFAELLHQCWEADAKKRPSEFKQIISILESMSNDTSLP 270  
OY 241 DKCNSFL 247  
DB 271 DKCNSFL 277

RESULT 4  
US-09-221-236-5  
; Sequence 5; Application US/09221236  
; Patent No. 6146841  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,236  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-236-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRAKWSQDKEVAVKKLKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60  
DB 31 YRAKWSQDKEVAVKKLKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90  
OY 61 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120  
DB 91 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150  
OY 121 DFGASRFHNHTTHMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLE 180  
DB 151 DFGASRFHNHTTHMSLVGTFPMMAPEVIOQLPVSETCDTYSYGVLMEMLTREVPFKGLE 210  
OY 181 GLQVAMLVVEKNERLTIPSSCPSPFAELLHQCWEADAKKRPSEFKQIISILESMSNDTSLP 240  
DB 211 GLQVAMLVVEKNERLTIPSSCPSPFAELLHQCWEADAKKRPSEFKQIISILESMSNDTSLP 270  
OY 241 DKCNSFL 247  
DB 271 DKCNSFL 277

RESULT 5  
US-09-221-416-5  
; Sequence 5; Application US/09221416  
; Patent No. 6153417  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,416  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-221-416-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLGS 60  
DB 31 YRAKWSODKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLGS 90  
QY 61 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120  
DB 91 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150  
QY 121 DFGASRFHNHTTHMSLVGTFPWWMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 180  
DB 151 DFGASRFHNHTTHMSLVGTFPWWMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 210  
QY 181 GLOVAVLVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSFKQIISILESMSNDTSLP 240  
DB 211 GLOVAVLVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSFKQIISILESMSNDTSLP 270  
QY 241 DKNSFL 247  
DB 271 DKNSFL 277

RESULT 6  
US-09-221-245-5  
Sequence 5, Application US/09221245  
Patent No. 6180358  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: MNI-050  
CURRENT APPLICATION NUMBER: US/09/221,245  
CURRENT FILING DATE: 1998-12-28  
EARLIER APPLICATION NUMBER: US 09/163,115  
EARLIER FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
-09-221-245-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLGS 60  
DB 31 YRAKWSODKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLGS 90  
QY 61 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120  
DB 91 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150  
QY 121 DFGASRFHNHTTHMSLVGTFPWWMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 180  
DB 151 DFGASRFHNHTTHMSLVGTFPWWMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 210  
QY 181 GLOVAVLVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSFKQIISILESMSNDTSLP 240  
DB 211 GLOVAVLVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSFKQIISILESMSNDTSLP 270  
QY 241 DKNSFL 247  
DB 271 DKNSFL 277

RESULT 7  
US-09-163-115-5  
Sequence 5, Application US/09163115A  
Patent No. 6183962  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: MNI-050  
CURRENT APPLICATION NUMBER: US/09/163,115A  
CURRENT FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-163-115-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLGS 60  
DB 31 YRAKWSODKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLGS 90  
QY 61 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120  
DB 91 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150  
QY 121 DFGASRFHNHTTHMSLVGTFPWWMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 180  
DB 151 DFGASRFHNHTTHMSLVGTFPWWMAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLE 210  
QY 181 GLOVAVLVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSFKQIISILESMSNDTSLP 240  
DB 211 GLOVAVLVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSFKQIISILESMSNDTSLP 270  
QY 241 DKNSFL 247  
DB 271 DKNSFL 277

RESULT 8  
US-09-221-528-5  
Sequence 5, Application US/09221528  
Patent No. 6190874  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: MNI-050  
CURRENT APPLICATION NUMBER: US/09/221,528  
CURRENT FILING DATE: 1998-12-28  
EARLIER APPLICATION NUMBER: 09/163,115  
EARLIER FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-221-528-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLGS 60  
DB 31 YRAKWSODKEVAVKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLGS 90  
QY 61 YDYINSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120



|||||  
Db 91 YDYNISNRSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150  
QY 121 DFGASRFHNHTTHMSLVGTFFPMAPEVIQSLPVSETCDTYSYGVILWEMLTREVPFKGLE 180  
Db 151 DFGASRFHNHTTHMSLVGTFFPMAPEVIQSLPVSETCDTYSYGVILWEMLTREVPFKGLE 210  
QY 181 GLOVAMLVVEKNERLTIPSSCPRSFAELLHQWEADAKKRPSEFKQIISILESMSNDTSLP 240  
Db 211 GLOVAMLVVEKNERLTIPSSCPRSFAELLHQWEADAKKRPSEFKQIISILESMSNDTSLP 270  
QY 241 DKNSFL 247  
Db 271 DKNSFL 277

## RESULT 9

US-09-593-553-5  
; Sequence 5, Application US/09593553

Patent No. 6200770

GENERAL INFORMATION:

APPLICANT: Acton, Susan

TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

FILE REFERENCE: MNI-050

CURRENT APPLICATION NUMBER: US/09/593,553

CURRENT FILING DATE: 2000-06-14

PRIOR APPLICATION NUMBER: 09/163,115

PRIOR FILING DATE: 1998-09-28

NUMBER OF SEQ ID NOS: 15

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 5

LENGTH: 455

TYPE: PRT

ORGANISM: Homo sapiens

US-09-593-553-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKMISQDKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60  
Db 31 YRAKMISQDKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90  
QY 61 YDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120  
Db 91 YDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150  
QY 121 DFGASRFHNHTTHMSLVGTFFPMAPEVIQSLPVSETCDTYSYGVILWEMLTREVPFKGLE 180  
Db 151 DFGASRFHNHTTHMSLVGTFFPMAPEVIQSLPVSETCDTYSYGVILWEMLTREVPFKGLE 210  
QY 181 GLOVAMLVVEKNERLTIPSSCPRSFAELLHQWEADAKKRPSEFKQIISILESMSNDTSLP 240  
Db 211 GLOVAMLVVEKNERLTIPSSCPRSFAELLHQWEADAKKRPSEFKQIISILESMSNDTSLP 270  
QY 241 DKNSFL 247  
Db 271 DKNSFL 277

## RESULT 10

US-09-221-237-5  
; Sequence 5, Application US/09221237

Patent No. 6214597

GENERAL INFORMATION:

APPLICANT: Acton, Susan

TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

FILE REFERENCE: MNI-050

CURRENT APPLICATION NUMBER: US/09/221,237

CURRENT FILING DATE: 1998-12-28

EARLIER APPLICATION NUMBER: 09/163,115

EARLIER FILING DATE: 1998-09-29

NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-237-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKMISQDKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60  
Db 31 YRAKMISQDKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90  
QY 61 YDYNISNRSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 120  
Db 91 YDYNISNRSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKIC 150  
QY 121 DFGASRFHNHTTHMSLVGTFFPMAPEVIQSLPVSETCDTYSYGVILWEMLTREVPFKGLE 180  
Db 151 DFGASRFHNHTTHMSLVGTFFPMAPEVIQSLPVSETCDTYSYGVILWEMLTREVPFKGLE 210  
QY 181 GLOVAMLVVEKNERLTIPSSCPRSFAELLHQWEADAKKRPSEFKQIISILESMSNDTSLP 240  
Db 211 GLOVAMLVVEKNERLTIPSSCPRSFAELLHQWEADAKKRPSEFKQIISILESMSNDTSLP 270  
QY 241 DKNSFL 247  
Db 271 DKNSFL 277

## RESULT 11

US-08-205-018-2  
; Sequence 2, Application US/08205018

Patent No. 5554523

GENERAL INFORMATION:

APPLICANT: Reddy, Usharani R.

TITLE OF INVENTION: No. 5554523el Protein kinase, Nucleic Acid

TITLE OF INVENTION: Sequences Encoding the Same and Methods Related

TITLE OF INVENTION: Thereto

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and

ADDRESSEE: No. 5554523ris

STREET: One Liberty Place - 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/205,018

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Gaumont, Rebecca R.

REGISTRATION NUMBER: 35,152

REFERENCE/DOCKET NUMBER: CH-0488

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 668 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-205-018-2

Query Match 37.7%; Score 490; DB 1; Length 668;  
Best Local Similarity 44.3%; Pred. No. 4.9e-43;  
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

QY 10 KEVAVKLLKIEKEAEI--LSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYN 67  
DB 147 EEVAVKKVRDL-KETDIKHLRKLKHPNITTFKGVCTQAPCYCLIMEFCAQGQLYEVLRA 205  
QY 68 RSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKICDFGASR- 126  
DB 206 RPTVPSL--LVDSMGIAAGMNYLHLH--KIHRDLKSPNMLITYDDVVKISDFGTSKE 260  
QY 127 FHNHTTHMSLVGTFFPMMAPEVIOSLPVSETCDTYSYGVVLWMLTREVPFKGLEQLQV 186  
DB 261 LSDKSTKMSFAGTVAMMAPEVIRNEPVSEKVDIWSFGVVLWMLLTGEIPYKDVDSALITW 320

187 LVEKNERLTIPSSCPSPFAELHQCWEADAKRPSFKQIISLESMSNDT-SLPDK 242  
321 GVGNSLHLVPSSCPDGFKILLRQCWNSKPRNRPFRQILLHLDIASADVLSTPQE 377

## RESULT 12

US-08-395-580-2  
Sequence 2, Application US/08395580  
Patent No. 5676945

## GENERAL INFORMATION:

APPLICANT: Usharani R. Reddy, David Pleasure and the Children's  
APPLICANT: Hospital of Philadelphia  
TITLE OF INVENTION: No. 5676945e1 Protein Kinase, Nucleic Acid  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Woodcock Washburn Kurtz Mackiewicz and No. 5676945e1s  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch disk, 720 KB  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/395,580  
FILING DATE: herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/205,018  
FILING DATE: 01-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Rebecca L. Ralph (formerly Gaumont)

REGISTRATION NUMBER: 35,152

REFERENCE/DOCKET NUMBER: CH-0488

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 859 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-395-580-2

Query Match 37.7%; Score 490; DB 1; Length 859;  
Best Local Similarity 44.3%; Pred. No. 7e-43;  
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

QY 10 KEVAVKLLKIEKEAEI--LSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYN 67

DB 147 EEVAVKKVRDL-KETDIKHLRKLKHPNITTFKGVCTQAPCYCLIMEFCAQGQLYEVLRA 205

QY 68 RSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKICDFGASR- 126  
DB 206 RPTVPSL--LVDSMGIAAGMNYLHLH--KIHRDLKSPNMLITYDDVVKISDFGTSKE 260

QY 127 FHNHTTHMSLVGTFFPMMAPEVIOSLPVSETCDTYSYGVVLWMLTREVPFKGLEQLQV 186  
DB 261 LSDKSTKMSFAGTVAMMAPEVIRNEPVSEKVDIWSFGVVLWMLLTGEIPYKDVDSALITW 320

QY 187 LVEKNERLTIPSSCPSPFAELHQCWEADAKRPSFKQIISLESMSNDT-SLPDK 242  
DB 321 GVGNSLHLVPSSCPDGFKILLRQCWNSKPRNRPFRQILLHLDIASADVLSTPQE 377

RESULT 13

PCT-US95-02792-2  
Sequence 2, Application PC/TUS9502792

GENERAL INFORMATION:

APPLICANT: Usharani R. Reddy, David Pleasure and the Children's  
APPLICANT: Hospital of Philadelphia  
TITLE OF INVENTION: Novel Protein Kinase, Nucleic Acid  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Woodcock Washburn Kurtz Mackiewicz and Norris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch disk, 720 KB  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/02792  
FILING DATE: herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/205,018  
FILING DATE: 01-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Rebecca L. Ralph (formerly Gaumont)

REGISTRATION NUMBER: 35,152

REFERENCE/DOCKET NUMBER: CH-0488

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 859 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US95-02792-2

Query Match 37.7%; Score 490; DB 5; Length 859;  
Best Local Similarity 44.3%; Pred. No. 7e-43;  
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

QY 10 KEVAVKLLKIEKEAEI--LSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYN 67  
DB 147 EEVAVKKVRDL-KETDIKHLRKLKHPNITTFKGVCTQAPCYCLIMEFCAQGQLYEVLRA 205

QY 68 RSEEMDMHMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLKICDFGASR- 126  
DB 206 RPTVPSL--LVDSMGIAAGMNYLHLH--KIHRDLKSPNMLITYDDVVKISDFGTSKE 260

QY 127 FHNHTTHMSLVGTFFPMMAPEVIOSLPVSETCDTYSYGVVLWMLTREVPFKGLEQLQV 186  
DB 321 GVGNSLHLVPSSCPDGFKILLRQCWNSKPRNRPFRQILLHLDIASADVLSTPQE 377

Db 261 LSDKSTKMSFAGTYVAMMAPEVIRNEPVSKEVDIWSFGVLTWELLTGEIPIYKDVDSSALITW 320

QY 187 LVEKNERLTTPSSCPSPFAELLHQCEADAKRPSFKOIIISLESMSNDT-SLPDK 242

Db 321 GVGSGNSLHLPVPSSCPDGFKILLRQCMNSKRRNRPSPRQILLHDIASADVLSTPOE 377

RESULT 14  
US-09-035

US-09-035-706-5  
Sequence 5: Application US/09035706

; Patent NO. 8001822  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dedhar, Shoukat  
 ; APPLICANT: Hannigan, Greg  
 ; TITLE OF INVENTION: Integrin-Linked Kinase and  
 ; TITLE OF INVENTION: its Uses  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Bozicevic & Reed, LLP  
 ; STREET: 285 Hamilton Avenue, Suite 200  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Sherwood, Pamela J  
 ; REGISTRATION NUMBER: 36,677  
 ; REFERENCE/DOCKET NUMBER: KLN-2C1P1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650-327-3400  
 ; TELEFAX: 650 327-3231  
 ; TELEX:  
 ;

```

! INFORMATION FOR SEQ ID NO: 5:
!
! SEQUENCE CHARACTERISTICS:
!     LENGTH: 263 amino acids
!     TYPE: amino acid
!     STRANDEDNESS: single
!     TOPOLOGY: linear
! MOLECULE TYPE: protein

```

MOLECULE TYPE: protein  
'  
US-09-035-706-5

|                          |        |                    |            |             |
|--------------------------|--------|--------------------|------------|-------------|
| Query Match              | 34.0%; | Score 442;         | DB 3;      | Length 263; |
| Best Local Similarity    | 40.2%; | Pred. No. 1.4e-38; |            |             |
| Matches 99; Conservative | 41;    | Mismatches 90;     | Indels 16; | Gaps 5;     |

```

OY      1 YRAKWI$QDEKAVAKKLKIE-----KEAELISVSHRNI IQFYGVILEPPNYGI    50
        :|:|:| | | | | | :|:|:| | | | | | :|:|:| | | | | |
Db      21 HRAEMHGSD--VAVKILMEQDFHAERVNEFLREVAIMKRLEHBNIVLFMGAVTQPENISI    78
OY      51 VTEYASLGSLYDYN-SNRSEEMDMHMTWATDVAKGMHYLHEAPVKVIHRDLKSBNV   109
        ||||| | | | | | :|:|:| | | | | | :|:|:| | | | | |
Db      79 VTEYLSRGSLYRLHKSGAREQLDERRRLSMAYDAKGMNLYHNRP-PIVHRDLKSPNL   137
OY     110 VIAADGVLKICDFGASRFPHNT--THMSLVGTFPWMAPEVIO$LPVSETCDTYSYGVALW   167
        ::|:|:| | | | | | :|:|:| | | | | | :|:|:| | | | | |
Db     138 LVDKKYTVKVCDFGLSRLLKASTFLLSSKSAAGTPEMAPEVLRLDBSPNEKSDVYSFGVILW   197
OY     168 EMLTREVPFKGLEGLQVAMLVVEKNERNLTIPSSCPRSFAELLHQCEADAKRPSFKOI I   227
        |:|:|:| | | | | | :|:|:| | | | | | :|:|:| | | | | |
Db     198 ELATLOQFWGNLNPAQVVAAVGFCKCRLEIFPNINPQVAAIIEGCWTNBPWKRPSPFATIM   257

```

|    |     |        |     |
|----|-----|--------|-----|
| QY | 228 | SILESM | 233 |
|    |     | : :    |     |
| Db | 258 | DLRPL  | 263 |

RESULT 15  
US-08-955

US-08-955-841-5  
Sequence 5. Application US/08955841

```

: Patent No. 6013782
:
: GENERAL INFORMATION:
:
: APPLICANT: Dedhar, Shoukat
:
: APPLICANT: Hannigan, Greg
:
: TITLE OF INVENTION: Integrin-Linked Kinase and
:
: TITLE OF INVENTION: its Uses
:
: NUMBER OF SEQUENCES: 11
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Bozicevic & Reed, LLP
:
: STREET: 285 Hamilton Avenue, Suite 200
:
: CITY: Palo Alto
:
: STATE: CA
:
: COUNTRY: USA

```

```

1  ZLF: 95504
2  COMPUTER READABLE FORM:
3  MEDIUM TYPE: Diskette
4  COMPUTER: IBM Compatible
5  OPERATING SYSTEM: DOS
6  SOFTWARE: FastSEO for Windows Version 2.0
7  CURRENT APPLICATION DATA:
8  APPLICATION NUMBER: US/08/955,841
9

```

FILING DATE: 800  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sherwood, Pamela J  
 REGISTRATION NUMBER: 36,677  
 REFERENCE/DOCKET NUMBER: KIN-2C1P1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-327-3400  
 TELEFAX: 650 327-3321  
 TELEX:

```

?      LENGTH: 263
?      INFORMATION FOR SEQ ID NO: 5:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 263 amino acids
?      TYPE: amino acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      MOLECULE TYPE: protein
?
US-08-955-841-5

```

MOLECULE 1  
US-08-955-841-5

|                       |                  |                    |            |             |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match           | 34.0%;           | Score 442;         | DB 3;      | Length 263; |
| Best Local Similarity | 40.2%;           | Pred. No. 1.4e-38; |            |             |
| Matches 99;           | Conservative 41; | Mismatches 90;     | Indels 16; | Gaps 5;     |

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | YRAKWTISQDKEVAVVKKLKIE-----KEAEILSVLSHRNITQFYGVILPEPNYGI       | 50  |
| Db | 21  | HRAEWGSD--VAVKILMEQDTHAERVNEFLREVALMKRLRHPNIVLFMGAVTQOPNLISI   | 78  |
| QY | 51  | VTEYASIGSLYDIYN-SNRSEEMDMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNV   | 109 |
| Db | 79  | VTEYLSRGSYLRLHLKSGAREQIDERRRLSMAYDVAKGMNYLHNRP-PIVHRDLKSNL     | 137 |
| QY | 110 | VIAADGVLKICDFGASRFHNHT--THMSLVGTFPWMAPEVIOQLPVSETCDTYSYGVILM   | 167 |
| Db | 138 | LVDKKTYTVKVCDFGLSRLKASTFLSSKSAAGTPEWMAPEVILDEPSENEKSDVYSFGVILM | 197 |
| QY | 168 | EMLTREVPFEGKLEGLQVAWLVEKNERLTISSCPRSFALILHQWEADAKRPSFKQII      | 227 |
| Db | 198 | ELATLQQPWNINLPAQVAAVGFCKRLEIPRLNLPQVAAITGCGWTNEPWKRPSFATIM     | 257 |
| QY | 228 | SILESM 233   |     |

Sun May 4 10:01:21 2003

Db 258 DLRRPL 263

Search completed: May 1, 2003, 20:39:41  
Job time : 28.5356 secs

us-09-757-982-5\_copy\_31\_277.ral





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:31:42 ; Search time 0.899582 Seconds

(without alignments)  
1602.986 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_407\_421

Perfect score: 74

Sequence: 1 KQNSKTTSKRGRGK 15

Scoring table: BLOSUM62

Searched: 283224 seqs, 96134422 residues

al number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73:\*

1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 46    | 62.2        | 216    | 2 D75567 | GTP cyclohydrolase  |
| 2          | 46    | 62.2        | 530    | 2 F86467 | hypothetical prote  |
| 3          | 45    | 60.8        | 857    | 2 A42861 | protein kinase TTK  |
| 4          | 44    | 59.5        | 425    | 2 S58672 | extra sex combs pr  |
| 5          | 43    | 58.1        | 683    | 2 T34103 | hypothetical prote  |
| 6          | 42    | 56.8        | 292    | 2 G88448 | protein C45G9.4 [i  |
| 7          | 42    | 56.8        | 294    | 2 E88448 | protein C45G9.9 [i  |
| 8          | 42    | 56.8        | 312    | 2 S56289 | regulatory protein  |
| 9          | 42    | 56.8        | 374    | 2 T19910 | hypothetical prote  |
| 10         | 42    | 56.8        | 752    | 2 S64750 | probable ATP-depen  |
| 11         | 42    | 56.8        | 1946   | 2 JC6032 | lactocepin (EC 3.4  |
| 12         | 41    | 55.4        | 453    | 2 G96695 | hypothetical prote  |
| 13         | 41    | 55.4        | 895    | 2 T34308 | hypothetical prote  |
| 14         | 40    | 54.1        | 208    | 2 S43434 | histone H1, testic  |
| 15         | 40    | 54.1        | 229    | 1 C45345 | vif protein - capr  |
| 16         | 40    | 54.1        | 298    | 2 S63238 | hypothetical prote  |
| 17         | 40    | 54.1        | 560    | 2 C71621 | exonuclease domain  |
| 18         | 40    | 54.1        | 1233   | 2 S56271 | hypothetical prote  |
| 19         | 39    | 52.7        | 81     | 2 A41949 | DNA topoisomerase   |
| 20         | 39    | 52.7        | 107    | 2 H84658 | hypothetical prote  |
| 21         | 39    | 52.7        | 319    | 2 AC3455 | UDP-galactose-4-epi |
| 22         | 39    | 52.7        | 401    | 2 T42655 | hypothetical prote  |
| 23         | 39    | 52.7        | 441    | 2 T25257 | nucleocapsid prote  |
| 24         | 39    | 52.7        | 546    | 2 T25257 | hypothetical prote  |
| 25         | 39    | 52.7        | 667    | 2 A40713 | cylicin I - bovine  |
| 26         | 39    | 52.7        | 673    | 1 BVECB  | excision nuclease   |
| 27         | 39    | 52.7        | 673    | 2 A99736 | DNA repair, excisi  |
| 28         | 39    | 52.7        | 673    | 2 B85586 | DNA repair, excisi  |
| 29         | 39    | 52.7        | 769    | 2 F89870 | serine proteinase   |

|    |      |      |      |          |                    |
|----|------|------|------|----------|--------------------|
| 30 | 39   | 52.7 | 905  | 2 A54654 | centromere protein |
| 31 | 39   | 52.7 | 952  | 2 S64473 | translation initia |
| 32 | 39   | 52.7 | 1481 | 1 QZDOP3 | pyrimidine synthes |
| 33 | 39   | 52.7 | 1485 | 1 ISZPT2 | DNA topoisomerase  |
| 34 | 39   | 52.7 | 1791 | 2 T02345 | hypothetical prote |
| 35 | 39   | 52.7 | 1898 | 2 T42440 | phospholipase C ho |
| 36 | 39   | 52.7 | 1922 | 2 T21581 | hypothetical prote |
| 37 | 39   | 52.7 | 3869 | 2 A48205 | All-1 protein +GTE |
| 38 | 38.5 | 52.0 | 1079 | 2 T38913 | translation initia |
| 39 | 38   | 51.4 | 80   | 2 F84090 | hypothetical prote |
| 40 | 38   | 51.4 | 142  | 2 I51651 | midline homolog -  |
| 41 | 38   | 51.4 | 142  | 2 JC4272 | pleiotrophic facto |
| 42 | 38   | 51.4 | 142  | 2 JC4168 | midline precursor  |
| 43 | 38   | 51.4 | 142  | 2 JC4273 | pleiotrophic facto |
| 44 | 38   | 51.4 | 258  | 2 S69056 | histone H1 - yeast |
| 45 | 38   | 51.4 | 293  | 2 B89802 | conserved hypothe  |

## ALIGNMENTS

## RESULT 1

D75567

GTP cyclohydrolase I - Deinococcus radiodurans (strain R1)

C/Species: Deinococcus radiodurans

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000

C/Accession: D75567

R/White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.,  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; M.

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999

A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A/Reference number: A75250; MUID:20036896; PMID:10567266

A/Accession: D75567

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-216 <WHI>

A/Cross-references: GB:AE001867; GB:AE000513; NID:g6457693; PIDN:AAF09628.1; PID:g64576  
A/Experimental source: strain R1

A/Genetics:

A/Gene: DR0036

A/Map position: 1

C/Superfamily: GTP cyclohydrolase I

Query Match 62.2%; Score 46; DB 2; Length 216;  
Best Local Similarity 76.9%; Pred. No. 3.5;  
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KQNSKTTSKRGRG 13

Db 181 KQNSSTTSAMRG 193

## RESULT 2

F86467

hypothetical protein F7P12.5 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001

C/Accession: F86467

R/Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marzall,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: F86467

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-530 <STO>  
A/Cross-references: GB:AE005172; NID:g10092384; PIDN:AAG12791.1; GSPDB:GN00141  
C/Genetics:  
A/Map position: 1

Query Match 62.2%; Score 46; DB 2; Length 530;  
Best Local Similarity 53.3%; Pred. No. 7.4;  
Matches 8; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 1 KONSSTTSKRGRKK 15  
:|||||:|||||  
Db 460 EQDTNKTTLKQGRK 474

## RESULT 3

A42861  
protein kinase TTK (EC 2.7.1.1) - human  
N/Alternate names: phosphotyrosine picked threonine kinase (PYT)  
C/Species: Homo sapiens (man)

Date: 10-Jun-1993 #sequence revision 18-Nov-1994 #text\_change 03-Nov-2000

Accession: A42861; S27971; I38144

Mills, G.B.; Schmandt, R.; McGill, M.; Amendola, A.; Hill, M.; Jacobs, K.; May, C.; Ro  
J. Biol. Chem. 267, 16000-16006, 1992

A/Title: Expression of TTK, a novel human protein kinase, is associated with cell prolif

A/Reference number: A42861; MUID:92348472; PMID:1639825

A/Accession: A42861

A/Molecule type: mRNA

A/Residues: 1-857 <MILL>

A/Cross-references: EMBL:M86699; NID:g340010

A/Note: sequence extracted from NCBI backbone (NCBIN:109875, NCBI:P.109876)

A/Note: It is uncertain whether Met-1 or Met-17 is the initiator

R/Mills, G.B.; Schmandt, R.; McGill, M.; Amendola, A.; Hill, M.; Jacobs, K.; May, C.; Ro  
submitted to the EMBL Data Library, February 1992

A/Description: Expression of TTK, a novel human protein kinase.

A/Reference number: S27971

A/Accession: S27971

A/Molecule type: mRNA

A/Residues: 17-857 <MILL>

A/Cross-references: EMBL:M86699; NID:g340010; PIDN:AAA61239.1; PID:g340011

R/Lindberg, R.A.; Fischer, W.H.; Hunter, T.

Oncogene 8, 351-359, 1993

A/Title: Characterization of a human protein threonine kinase isolated by screening an e

A/Reference number: I38144; MUID:93149596; PMID:7678926

A/Accession: I38144

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 'MCMR', 504-767, 'V', 769-802, 'GI' <LIN>

A/Cross-references: EMBL:X70500; NID:g312815; PIDN:CAA49912.1; PID:g312816

Genetics:

Gene: GDB:TTK

A/Cross-references: GDB:455142

A/Map position: 7p12-7cen

C/Superfamily: protein kinase homology

C/Keywords: phosphotransferase

F/523-791/Domain: protein kinase homology <KIN>

Query Match 60.8%; Score 45; DB 2; Length 857;  
Best Local Similarity 69.2%; Pred. No. 16;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 NSSKTSKRGRKK 15  
:|||||:|||||  
Db 845 SSSKTFEKKRGRKK 857

## RESULT 4

S58672  
extra sex combs protein - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 26-May-2000

C/Accession: S58672

R/Gutjahr, T.; Frei, E.; Spicer, C.; Baumgartner, S.; White, R.A.H.; Noll, M.

EMBO J. 14, 4296-4306, 1995

A/Title: The polycomb-group gene, extra sex combs, encodes a nuclear member of the WD-4

A/Reference number: S58672; MUID:96016202; PMID:7556071

A/Accession: S58672

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: DNA

A/Residues: 1-425 <GUT>

A/Cross-references: GB:L41867; NID:g1050996; PIDN:AAA86427.1; PID:g1050997

C/Genetics:

A/Gene: FlyBase:esc

A/Cross-references: FlyBase:FBgn000588

C/Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 59.5%; Score 44; DB 2; Length 425;  
Best Local Similarity 60.0%; Pred. No. 13;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 KONSSTTSKRGRKK 15  
:|||||:|||||  
Db 36 KSPSSSTRSKRRGR 50

## RESULT 5

T34103  
hypothetical protein C17G10.6 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C/Accession: T34103

R/Johnson, D.

submitted to the EMBL Data Library, June 1995

A/Description: The sequence of C. elegans cosmid C17G10.

A/Reference number: Z21476

A/Accession: T34103

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-683 <JOH>

A/Cross-references: EMBL:U28739; PIDN:AAB93453.1; GSPDB:GN00020; CESP:C17G10.6

A/Experimental source: strain Bristol N2; clone C17G10

C/Genetics:

A/Gene: CESP:C17G10.6

A/Map position: 2

A/Introns: 37/1; 129/2; 156/2; 178/2; 258/3; 291/3; 521/3; 611/1

Query Match 58.1%; Score 43; DB 2; Length 683;  
Best Local Similarity 53.3%; Pred. No. 28;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 KONSSTTSKRGRKK 15  
:|||||:|||||  
Db 557 KKKSKKNNKRRGRKK 571

## RESULT 6

G88448  
protein C45G9.4 [imported] - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

C/Accession: G88448

R/anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo

A/Reference number: A75000; MUID:99069613; PMID:9851916

A/Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.banger.ac.uk/Projects/C\_el

A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an

A/Accession: G88448

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-292 <STO>

A/Cross-references: GB:chr\_III; PIDN:AAA62556.1; PID:g687878; GSPDB:GN00021; CESP:C45G9

C/Genetics:

A/Gene: C45G9.4

A/Map position: 3

Query Match 56.8%; Score 42; DB 2; Length 292;

Best Local Similarity 57.1%; Pred. No. 20;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 QNSSKTSKRGGK 15  
Db 30 RRSSTSKRKRK 43

## RESULT 7

E88448

protein C45G9.9 [imported] - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans

C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C/Accession: E88448

R/Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998

A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A/Reference number: A75000; MUID:99069613; PMID:9851916

A/Note: See websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

Accession: E88448

Status: preliminary

A/Molecule type: DNA

A/Residues: 1-294 <STO>

A/Cross-references: GB:chr\_III; PIDN:AAA62552.1; PID:6687874; GSPDB:GN00021; CESP:C45G9.  
C/Genetics:

A/Gene: C45G9.9

A/Map position: 3

Query Match 56.8%; Score 42; DB 2; Length 294;  
Best Local Similarity 57.1%; Pred. No. 20;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

## RESULT 8

S56289

regulatory protein PHO4 - Yeast (Saccharomyces cerevisiae)

N/Alternate names: phosphate system positive regulatory protein; protein R007; protein Y  
C/Species: Saccharomyces cerevisiae

C/Date: 02-Sep-1995 #sequence\_revision 12-Apr-1996 #text\_change 17-Mar-1999  
C/Accession: S56289; A23482; S62245; S63839

R/Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasamura, S.I.; Sasamu  
submitted to the EMBL Data Library, May 1995

Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces ce  
Reference number: S56186

Accession: S56289

A/Molecule type: DNA

A/Residues: 1-312 <MUR>

A/Cross-references: EMBL:D50617; NID:g836685; PID:d1009914; PID:g836789; MIPS:YFR034c  
R/Legrain, M.; De Wilde, M.; Hilger, F.

Nucleic Acids Res. 14, 3059-3073, 1986

A/Title: Isolation, physical characterization and expression analysis of the Saccharomy  
A/Reference number: A23482; MUID:86176785; PMID:3008105

Accession: A23482

A/Molecule type: DNA

A/Residues: 1-289, 'RP', 292, 'RPPWRPAGTSVT', 302, 'S', 304, 'T' <LEG>

A/Cross-references: EMBL:X03719

R/Murakami, Y.  
submitted to the EMBL Data Library, December 1994

A/Reference number: S62230

A/Accession: S62245

A/Molecule type: DNA

A/Residues: 1-312 <MUR>

A/Cross-references: EMBL:D44602; NID:g893419; PID:d1008636; PID:g893429  
R/Eki, T.; Naitou, M.; Hagiwara, H.; Abe, M.; Ozawa, M.; Sasamura, S.I.; Sasamura, M.; T  
Yeast 12, 177-190, 1996

A/Title: Fifteen open reading frames in a 30.8 kb region of the right arm of chromosome  
A/Reference number: S63830; MUID:96287654; PMID:8686381

A/Accession: S63839

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-312 <EKI>

A/Cross-references: EMBL:D44602; NID:g893419; PID:d1008636; PID:g893429

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995

C/Genetics:

A/Gene: SGD:PHO4

A/Cross-references: SGD:S0001930; MIPS:YFR034c

A/Map position: 6R

## RESULT 9

T19910

hypothetical protein C43F9.6 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C/Accession: T19910

R/Mortimore, B.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z19195

A/Accession: T19910

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-374 <WIL>

A/Cross-references: EMBL:Z82262; PIDN:CAB05149.1; GSPDB:GN00022; CESP:C43F9.6

A/Experimental source: clone C43F9

C/Genetics:

A/Gene: CESP:C43F9.6

A/Map position: 4

A/Introns: 22/1; 82/2; 104/1; 151/1; 193/1; 235/2; 292/3; 335/3

Query Match 56.8%; Score 42; DB 2; Length 374;  
Best Local Similarity 69.2%; Pred. No. 25;  
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

## RESULT 10

S64750

probable ATP-dependent RNA helicase DRS1 - yeast (Saccharomyces cerevisiae)

N/Alternate names: protein L1345; protein YL1008w

C/Species: Saccharomyces cerevisiae

C/Date: 01-Aug-1995 #sequence\_revision 24-May-1996 #text\_change 02-Feb-2001  
C/Accession: S64750; S31248; S70567

R/Miosga, T.; Zimmermann, F.K.  
submitted to the Protein Sequence Database, May 1996

Accession: S64750

A/Molecule type: DNA

A/Residues: 1-752 <MIO>

A/Cross-references: EMBL:Z73113; NID:g1360170; PIDN:CA97452.1; PID:e245447; PID:g13601,  
A/Experimental source: strain S288C

R/Ripmaster, T.L.; Vaughn, G.P.; Woolford Jr., J.L.  
Proc. Natl. Acad. Sci. U.S.A. 89, 11131-11135, 1992

A/Title: A putative ATP-dependent RNA helicase involved in Saccharomyces cerevisiae ribo  
A/Reference number: S31248; MUID:93087480; PMID:1454790

Accession: S31248

A/Molecule type: DNA

A/Residues: 'MTK', 34, 'SRLRLRSGRVTR', 51, 'RLVR', 56, 'ITSMRMFMRTWT', 69-752 <RIP>

A/Cross-references: EMBL:L00683; NID:g171655; PIDN:AAA34666.1; PID:g171656

R/Miosga, T.; Zimmermann, F.K.  
Yeast 12, 693-708, 1996



A/Title: Sequence analysis of the GEN12 region of *Saccharomyces cerevisiae* on a 43.7 kb  
e conductance regulator protein CTR.  
A/Reference number: S70557; MUID:96405918; PMID:8810043  
A/Accession: S70567  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-752 <MW>  
A/Cross-references: EMBL:X91488; NID:g1495203; PIDN:CAA62783.1; PID:el99012; PID:g149521  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995  
C/Genetics:  
A/Gene: SGD:DRS1  
A/Cross-references: SGD:S0003931; MIPS:YLL008w  
A/Map position: 12L  
C/Keywords: ATP; nucleotide binding; P-loop  
F/275-282/Region: nucleotide-binding motif A (P-loop)  
F/381-386/Region: nucleotide-binding motif B  
F/385-388/Region: DEAD motif

Query Match 56.8%; Score 42; DB 2; Length 752;  
Best Local Similarity 57.1%; Pred. No. 44;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGK 14  
| : ||| ||| : ||  
Db 34 KVEAKKTTKRRGK 47

RESULT 11  
JC6032  
lactocepin (EC 3.4.21.96) precursor [similarity] - *Lactobacillus delbrueckii* subsp. bulg  
N/Alternate names: cell envelope-associated serine proteinase prtp  
C/Species: *Lactobacillus delbrueckii* subsp. bulgaricus  
C/Date: 31-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 21-Jul-2000  
C/Accession: JC6032  
R/Gilbert, C.; Atlan, D.; Blanc, B.; Portalier, R.; Germond, J.E.; Lapierre, L.; Mollet,  
J. Bacteriol. 178, 3059-3065, 1996  
A/Title: A new cell surface proteinase: Sequencing and analysis of the prtB gene from *La*  
A/Reference number: JC6032; MUID:96236017; PMID:8655480  
A/Accession: JC6032  
A/Molecule type: DNA  
A/Residues: 1-1946 <GIL>  
A/Cross-references: GB:L48487; NID:g6013471; PIDN:AAC41529.1; PID:g1381114  
A/Experimental source: NCD01489  
A/Note: neither the complete nucleic acid sequence nor the complete translation are show  
C/Genetics:  
A/Gene: prtB  
C/Superfamily: lactocepin; subtilisin homology  
C/Keywords: hydrolase; serine proteinase  
-34/Domain: signal sequence #status predicted <SIG>  
-5-1946/Product: cell surface proteinase #status predicted <MAT>  
F/213-631/Domain: subtilisin homology #status atypical <SBT>

Query Match 56.8%; Score 42; DB 2; Length 1946;  
Best Local Similarity 60.0%; Pred. No. 98;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGK 15  
| : ||| ||| ||  
Db 1789 KKTDSKTSKRSACK 1803

RESULT 12  
G96695  
hypothetical protein F5A8.9 [imported] - *Arabidopsis thaliana*  
C/Species: *Arabidopsis thaliana* (mouse-ear cross)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: G96695  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Hutzar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: G96695  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-453 <STO>  
A/Cross-references: GB:AE005173; NID:g4204282; PIDN:AAD10663.1; GSPDB:GN00141  
C/Genetics:  
A/Gene: F5A8.9  
A/Map position: 1

Query Match 55.4%; Score 41; DB 2; Length 453;  
Best Local Similarity 53.3%; Pred. No. 42;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGK 15  
| : ||| ||| : ||  
Db 164 KENSGVAESSRKGK 178

RESULT 13  
T34308  
hypothetical protein F55C12.5 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C/Accession: T34308  
R/Latreille, P.  
submitted to the EMBL Data Library, November 1995  
A/Description: The sequence of *C. elegans* cosmid F55C12.  
A/Reference number: Z21503  
A/Accession: T34308  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-895 <LAT>  
A/Cross-references: EMBL:U41107; PIDN:AAC71164.1; GSPDB:GN00020; CESP:F55C12.5  
A/Experimental source: strain Bristol N2; clone F55C12  
C/Genetics:  
A/Gene: CESP:F55C12.5  
A/Map position: 2  
A/Introns: 6/3; 25/3; 66/3; 240/3; 481/2; 535/3; 687/3; 791/3; 827/2

Query Match 55.4%; Score 41; DB 2; Length 895;  
Best Local Similarity 53.3%; Pred. No. 75;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGK 15  
| : ||| ||| : ||  
Db 752 KNNSEGTTEKTKGK 766

RESULT 14  
S43434  
histone H1, testicular - mouse  
C/Species: *Mus musculus* (house mouse)  
C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Jun-2000  
C/Accession: S43434; S33226  
R/Drabant, B.; Bode, C.; Doenecke, D.  
Biochim. Biophys. Acta 1216, 311-313, 1993  
A/Title: Structure and expression of the mouse testicular H1 histone gene (H1c).  
A/Reference number: S43434; MUID:94060108; PMID:8241275  
A/Accession: S43434  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-208 <DRA>  
A/Cross-references: EMBL:X72805; NID:g1934959; PIDN:CAA51325.1; PID:g297754  
A/Note: the authors did not translate the codon for residue 1  
C/Superfamily: histone H1  
C/Keywords: DNA binding; nucleus

Query Match 54.1%; Score 40; DB 2; Length 208;

Best Local Similarity 80.0%; Pred. No. 32;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 KTTSKRRGKK 15  
|:|||||  
Db 19 KPSSKRRGKK 28

## RESULT 15

C45345

vif protein - caprine arthritis-encephalitis virus (strain CO)  
N;Alternate names: orf-Q protein; sor protein

C;Species: caprine arthritis-encephalitis virus, CAEV

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999

C;Accession: C45345

R;Salicrull, M.; Querat, G.; Konings, D.A.M.; Vigne, R.; Clements, J.E.  
Virology 179, 347-364, 1990

A;Title: Nucleotide sequence and transcriptional analysis of molecular clones of CAEV with  
Accession: A45345; MUID:91021037; PMID:2171210

Accession: C45345

Molecule type: mRNA

Residues: 1-229 &lt;SAL&gt;

A;Cross-references: GB:M33677; NID:G323294; PIDN:AAA91827.1; PID:G323297

C;Genetics:

A;Gene: vif

C;Superfamily: Viena virus vif protein

Query Match 54.1%; Score 40; DB 1; Length 229;  
Best Local Similarity 57.1%; Pred. No. 35;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 QNSSKTSKRRGKK 15  
|:|||||  
Db 2 QNSSRHOQKRRKK 15

Search completed: May 1, 2003, 20:38:42  
Job time: 3.89958 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:13:36 ; Search time 0.753138 Seconds

(without alignments)

826.070 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_407\_421

Perfect score: 74

Sequence: 1 KONSSTKSKRGKK 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

al number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 46    | 62.2        | 216    | 1     | GCH1_DEIRA  |
| 2          | 45    | 60.8        | 841    | 1     | TKK_HUMAN   |
| 3          | 42    | 56.8        | 292    | 1     | YQ14_CAEEL  |
| 4          | 42    | 56.8        | 294    | 1     | YQ19_CAEEL  |
| 5          | 42    | 56.8        | 312    | 1     | PHO4_YEAST  |
| 6          | 42    | 56.8        | 752    | 1     | DRS1_YEAST  |
| 7          | 40    | 54.1        | 207    | 1     | HIT_YEAST   |
| 8          | 40    | 54.1        | 229    | 1     | VIF_CAEVC   |
| 9          | 40    | 54.1        | 298    | 1     | YN05_YEAST  |
| 10         | 40    | 54.1        | 1233   | 1     | YF16_YEAST  |
| 11         | 39    | 52.7        | 667    | 1     | CYL1_BOVIN  |
| 12         | 39    | 52.7        | 673    | 1     | UVRB_ECOLI  |
| 13         | 39    | 52.7        | 906    | 1     | CENC_MOUSE  |
| 14         | 39    | 52.7        | 952    | 1     | IF41_YEAST  |
| 15         | 39    | 52.7        | 1485   | 1     | TOP2_SCHPO  |
| 16         | 39    | 52.7        | 2185   | 1     | PYR1_DICDI  |
| 17         | 39    | 52.7        | 3866   | 1     | HRX_MOUSE   |
| 18         | 38.5  | 52.0        | 1079   | 1     | IF2P_SCHPO  |
| 19         | 38    | 51.4        | 125    | 1     | H2B_ACRFO   |
| 20         | 38    | 51.4        | 142    | 1     | PTAI_XENLA  |
| 21         | 38    | 51.4        | 142    | 1     | PTA2_XENLA  |
| 22         | 38    | 51.4        | 258    | 1     | H1_YEAST    |
| 23         | 38    | 51.4        | 351    | 1     | COA2_POVBA  |
| 24         | 38    | 51.4        | 351    | 1     | COA2_POVBA  |
| 25         | 38    | 51.4        | 455    | 1     | ORC5_SCHPO  |
| 26         | 38    | 51.4        | 793    | 1     | YF06_MYCPN  |
| 27         | 38    | 51.4        | 794    | 1     | YB52_MYCPN  |
| 28         | 38    | 51.4        | 814    | 1     | TOP1_SCHPO  |
| 29         | 38    | 51.4        | 856    | 1     | ENV_FVISC   |
| 30         | 38    | 51.4        | 914    | 1     | IF42_YEAST  |
| 31         | 38    | 51.4        | 1711   | 1     | CHD1_MOUSE  |
| 32         | 38    | 51.4        | 1920   | 1     | PCNT_MOUSE  |
| 33         | 37    | 50.0        | 327    | 1     | KAPR_DICDI  |

|    |    |      |      |   |            |
|----|----|------|------|---|------------|
| 34 | 37 | 50.0 | 339  | 1 | RL29_SPICI |
| 35 | 37 | 50.0 | 376  | 1 | GUNK_FUSOX |
| 36 | 37 | 50.0 | 433  | 1 | ELT2_CAEEL |
| 37 | 37 | 50.0 | 471  | 1 | RB97_DROME |
| 38 | 37 | 50.0 | 482  | 1 | Y138_METJA |
| 39 | 37 | 50.0 | 561  | 1 | SR72_SCHPO |
| 40 | 37 | 50.0 | 882  | 1 | RA50_PYRFU |
| 41 | 37 | 50.0 | 892  | 1 | ATX7_HUMAN |
| 42 | 37 | 50.0 | 914  | 1 | PBPA_BACSV |
| 43 | 37 | 50.0 | 939  | 1 | XPC_HUMAN  |
| 44 | 37 | 50.0 | 1093 | 1 | SW14_YEAST |
| 45 | 37 | 50.0 | 1123 | 1 | RBM6_HUMAN |

## ALIGNMENTS

## RESULT 1

GCH1\_DEIRA

ID GCH1\_DEIRA

AC Q9RYB4

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE GTP cyclohydrolase I (EC 3.5.4.16) (GTP-CH-I).

GN FOLE OR DR0036.

OS Deinococcus radiodurans.

OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;

OC Deinococcaceae; Deinococcus.

OX NCBI\_TaxID=1299;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=R1;

RX MEDLINE=20036896; PubMed=10567266;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,

RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,

RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,

RA Makarova K.S., Aravind L., Daly M.J., Mintana W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,

RA Fraser C.M.;

RT "Genome sequence of the radioresistant bacterium Deinococcus

RT radiodurans R1.";

RL Science 286:1571-1577(1999).

CC -!- CATALYTIC ACTIVITY: GTP + 2 H(2)O = formate + 2-amino-4-hydroxy-6-

CC (erythro-1,2,3-trihydroxypropyl)dihydropteridine triphosphate.

CC -!- PATHWAY: Tetrahydrofolate biosynthesis; first step.

CC -!- SUBUNIT: HOMOPOLYMER (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE GTP CYCLOHYDROLASE I FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; AE001867; AAF09628.1; --

CC HSSP; P27511; 1GTP.

CC TIGR; DR0036; --

CC InterPro; IPR001474; GTP\_cyclohydrol.

CC Pfam; PF01227; GTP\_cyclohydrol; 1.

CC ProDom; PD003330; GTP\_cyclohydrol; 1.

CC TIGRfam; TIGR0063; FOLE; 1.

CC PROSITE; PS00859; GTP\_CYCLOHYDROL\_1; 1.

CC PROSITE; PS00860; GTP\_CYCLOHYDROL\_2; 1.

CC One-carbon metabolism; Hydrolase; Complete proteome.

CC DISULFID 102 173 BY SIMILARITY.

CC SEQUENCE 216 AA; 23799 MW; 149F5281A60D4C4C CRC64;

Query Match. 62.2%; Score 46; DB 1; Length 216;  
Best Local Similarity 76.9%; Pred. No. 0.88;



Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRG 13  
Db 181 KONSSTTSAMRG 193

## RESULT 2

TTK\_HUMAN STANDARD; PRT; 841 AA.  
AC P33981;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Dual specificity protein kinase TTK (EC 2.7.1.-) (PYT).  
GN TTK OR MPS1L1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OC NCBI\_TaxID=9606;

## [1] SEQUENCE FROM N.A.

RA MEDLINE=92348472; PubMed=1639825;  
RA Mills G.B., Schmandt R., McGill M., Amendola A., Hill M.,  
RA Jacobs K., May C., Rodricks A.-M., Campbell S., Hogg D.;  
RT "Expression of TTK, a novel human protein kinase, is associated with  
RT cell proliferation.";  
RL J. Biol. Chem. 267:16000-16006(1992).

## [2] SEQUENCE OF 509-776 FROM N.A.

RA MEDLINE=92065863; PubMed=1956325;  
RA Lindberg R.A., Hunter T.;  
RL Hanks S.K., Quinn A.M.;  
RL Meth. Enzymol. 200:38-62(1991).  
CC -!- FUNCTION: PHOSPHORYLATES PROTEINS ON SERINE, THREONINE, AND  
CC TYROSINE. PROBABLY ASSOCIATED WITH CELL PROLIFERATION.  
CC -!- TISSUE SPECIFICITY: PRESENT IN RAPIDLY PROLIFERATING CELL LINES.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M86699; AAA61239.1; -.  
PIR; S27971; S27971.  
HSSP; Q00534; 1B18.

DR Genew; HGNC:12401; TTK.  
DR MIM; 604092; -.

DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.

DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.

DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.

DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS0011; PROTEIN\_KINASE\_DOM; 1.

KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
KW Tyrosine-protein kinase.

FT DOMAIN 509 775 PROTEIN KINASE.  
FT NP\_BIND 515 523 ATP (BY SIMILARITY).

FT BINDING 537 537 ATP (BY SIMILARITY).  
FT ACT\_SITE 631 631 BY SIMILARITY.

FT CONFLICT 752 752 L -> V (IN REF. 2).  
SQ SEQUENCE 841 AA; 95279 MW; ACDEC2ACD15EF08C CRC64;

Query Match 60.8%; Score 45; DB 1; Length 841;  
Best Local Similarity 69.2%; Pred. No. 5.2;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 NSSKTTSKRRGK 15  
Db 829 SSSKTFEKKRGK 841

## RESULT 3

YQ14\_CAEEL STANDARD; PRT; 292 AA.  
ID YQ14\_CAEEL  
AC Q09505;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Hypothetical 32.4 kDa protein C45G9.4 in chromosome III.  
GN C45G9.4.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OC NCBI\_TaxID=6239;

## [1] SEQUENCE FROM N.A.

RA STRAIN=Bristol N2;  
RA Bentley D., Waterston R.;  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U21323; AAA62556.1; -.  
DR WormPep; C45G9.4; CE01855.  
KW Hypothetical protein.  
FT DOMAIN 9 84 LYS-RICH.  
SQ SEQUENCE 292 AA; 32454 MW; 589FCD425363A9B8 CRC64;

Query Match 56.8%; Score 42; DB 1; Length 292;  
Best Local Similarity 57.1%; Pred. No. 5.6;

Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 QNSKTTSKRRGK 15  
Db 30 RRSSTTSKRRGK 43

## RESULT 4

YQ19\_CAEEL STANDARD; PRT; 294 AA.  
ID YQ19\_CAEEL  
AC Q09507;

DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Hypothetical 32.7 kDa protein C45G9.9 in chromosome III.  
GN C45G9.9.

OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.  
OC NCBI\_TaxID=6239;

RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;  
RA Bentley D., Waterston R.;

RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch).

CC -----  
CC EMBL; U21323; AAA62552.1; -  
DR WormPep; C45G9.9; CE01860.  
KW Hypothetical protein.  
FT DOMAIN 9 88 LYS-RICH.  
SQ SEQUENCE 294 AA; 32738 MW; 856B595C1FCC5448 CRC64;

Query Match 56.8%; Score 42; DB 1; Length 294;  
Best Local Similarity 57.1%; Pred. No. 5.7;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 QNSSKTSKRGRGK 15  
Db 30 RRSKTSKRGRGK 43

RESULT 5  
PHO4\_YEAST  
PHO4\_YEAST STANDARD; PRT; 312 AA.  
P07270;  
01-APR-1988 (Rel. 07, Created)  
01-FEB-1991 (Rel. 17, Last sequence update)  
15-JUN-2002 (Rel. 41, Last annotation update)  
DE Phosphate system positive regulatory protein PHO4.  
GN PHO4 OR YPR034C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86176785; PubMed=3008105;  
RA Legrain M., de Wilde M., Hilger F.;  
RT "Isolation, physical characterization and expression analysis of the  
RL Saccharomyces cerevisiae positive regulatory gene PHO4.";  
Nucleic Acids Res. 14:3059-3073(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89364688; PubMed=2505053;  
RA Yoshida K., Kuromitsu Z., Ogawa N., Oshima Y.;  
RT "Mode of expression of the positive regulatory genes PHO2 and PHO4 of  
RL the phosphatase regulon in Saccharomyces cerevisiae.";  
Mol. Gen. Genet. 217:31-39(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
MEDLINE=95400292; PubMed=7670463;  
Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,  
Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,  
Yamazaki M., Tashiro H., Eki T.;  
RT "Analysis of the nucleotide sequence of chromosome VI from  
RL Saccharomyces cerevisiae.";  
Nat. Genet. 10:261-268(1995).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
MEDLINE=96287654; PubMed=8686381;  
Eki T., Naitou M., Hagiwara H., Abe M., Ozawa M., Sasanuma S.-I.,  
Sasanuma M., Tsuchiya Y., Shibata T., Watanabe K., Ono A.,  
Yamazaki M.-A., Tashiro H., Hanaoka F., Murakami Y.;  
RT "Fifteen open reading frames in a 30.8 kb region of the right arm of  
RL chromosome VI from Saccharomyces cerevisiae.";  
Yeast 12:177-190(1996).  
RN [5]  
RP DOMAINS.  
RX MEDLINE=90220608; PubMed=2183025;  
Ogawa N., Oshima Y.;  
RT "Functional domains of a positive regulatory protein, PHO4, for  
RT transcriptional control of the phosphatase regulon in Saccharomyces  
RL cerevisiae.";  
Mol. Cell. Biol. 10:2224-2236(1990).  
DT [6]

RP HELIX-LOOP-HELIX MOTIF.  
RX MEDLINE=91021495; PubMed=2220078;  
RA Berben G., Legrain M., Gilliquet V., Hilger F.;  
RT "The yeast regulatory gene PHO4 encodes a helix-loop-helix motif.";  
RL Yeast 6:451-454(1990).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 251-312.  
RX MEDLINE=97447573; PubMed=9303313;  
RA Shimizu T., Tomoto A., Ihara K., Shimizu M., Kyogoku Y., Ogawa N.,  
RA Oshima Y., Hakoshima T.;  
RT "Crystal structure of PHO4 bHLH domain-DNA complex: flanking base  
RT recognition.";  
EMBO J. 16:4689-4697(1997).  
RL  
CC -1- FUNCTION: THIS ACTIVATOR REGULATES THE EXPRESSION OF REPRESSIBLE  
CC PHOSPHATASE AT THE TRANSCRIPTIONAL LEVEL. PHO4 IS PRESUMED TO  
CC INTERACT WITH THE UAS (UPSTREAM ACTIVATING SEQUENCE) OF SEVERAL  
CC PHOSPHATASE ENCODING PHO GENES.  
CC -1- SUBUNIT: BINDS DNA AS A DIMER OR HIGHER COMPLEX.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- PTM: PHOSPHORYLATED BY PHO85 WHEN INORGANIC PHOSPHATE IS PRESENT.  
CC THIS PREVENTS PHO4 FROM ACTIVATING THE STRUCTURAL PHO GENES.  
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF  
CC TRANSCRIPTION FACTORS.  
CC -1- SIMILARITY: TWO REGIONS OF STRONG HOMOLOGY TO PHO2 ARE FOUND.  
CC ALSO SIMILAR TO N. CRASSA NUC1.  
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN IN THE C-TERMINUS  
CC DUE TO A FRAMESHIFT.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>  
CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
-----  
DR EMBL; X03719; CAA27345.1; ALT\_FRAME.  
DR EMBL; D50617; BAA09273.1; -  
DR PIR; A23482; A23482.  
DR PIR; 1A0A; 18-MAR-98.  
DR TRANSFAC; T00690; -  
DR SGD; S0001930; PHO4.  
DR InterPro; IPR001092; HLH\_basic.  
DR Pfam; PF00010; HLH; 1.  
DR SMART; SM00353; HLH; 1.  
DR PROSITE; PS00038; HLH\_1; 1.  
DR PROSITE; PS50888; HLH\_2; 1.  
KW Transcription regulation; Activator; DNA-binding; Phosphorylation;  
KW Nuclear protein; 3D-structure.  
FT DOMAIN 1 109 INVOLVED IN TRANSCRIPTIONAL ACTIVATION.  
FT DOMAIN 203 227 INVOLVED IN OLIGOMERIZATION.  
FT DNA\_BIND 251 ? BASIC DOMAIN (POTENTIAL).  
FT DOMAIN ? 307 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).  
FT CONFLICT 269 269 P -> A (IN REF. 3 AND 4).  
FT CONFLICT 310 310 G -> V (IN REF. 3 AND 4).  
SQ SEQUENCE 312 AA; 34072 MW; 6DB926A998A09241 CRC64;

Query Match 56.8%; Score 42; DB 1; Length 312;  
Best Local Similarity 64.3%; Pred. No. 6;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 QNSSKTSKRGRGK 14  
Db 148 KSNSSPYLNKRGRGK 161

RESULT 6  
DRS1\_YEAST  
ID DRS1\_YEAST STANDARD; PRT; 752 AA.  
AC P32892;  
DT 01-OCT-1993 (Rel. 27, Created).  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable ATP-dependent RNA helicase DRS1.  
GN DRS1 OR YLL008W OR L1345.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93087480; PubMed=1454790;  
RA Ripmaster T.L., Vaughn G.P., Woolford J.L. Jr.;  
RT "A putative ATP-dependent RNA helicase involved in Saccharomyces  
cerevisiae ribosome assembly."; Proc. Natl. Acad. Sci. U.S.A. 89:11131-11135(1992).  
RL [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=S288C / FY23;  
RX MEDLINE=96405918; PubMed=8810043;  
RA Miosga T., Zimmermann F.K.;  
RT "Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on  
a 43.7 kb fragment of chromosome XII including an open reading frame  
homologous to the human cystic fibrosis transmembrane conductance  
regulator protein CFTR."; yeast 12:693-708(1996).  
RL -1- FUNCTION: PROBABLE HELICASE INVOLVED IN RIBOSOME ASSEMBLY.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L00683; AAA34666.1; -  
DR EMBL; X91488; CAA62783.1; -  
DR EMBL; Z73113; CAA97452.1; -  
DR PIR; S31248; S31248.  
DR HSSP; Q58083; 1HV8.  
DR SGD; S0003931; DRS1.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR000629; DEAD\_box.  
DR InterPro; IPR001650; Helicase\_C.  
DR Pfam; PF00270; DEAD; 1.  
DR Pfam; PF00271; helicase\_C; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR SMART; SM00490; HELICC; 1.  
DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
ATP-binding; RNA-binding; Helicase; Nuclear protein.  
FT DOMAIN 170 190 POLY-GLU.  
FT NP BIND 275 282 ATP (POTENTIAL).  
FT SITE 385 388 DEAD BOX.  
FT CONFLICT 1 68 MVTGKKYKYNLDFVPTISDEDDVPIILDSDEKYEAKKTT  
KRRGKNNKKKVSSEGNLDEVDVHEDLD -> MTKKSRRLRL  
RRSGRVIRTRKRLVRGITSMMRMFMRTWT (IN REF. 1).  
SQ SEQUENCE 752 AA; 84843 MW; 60747607A6E5E4A8 CRC64;  
Query Match 56.8%; Score 42; DB 1; Length 752;  
Best Local Similarity 57.1%; Pred. No. 15;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
OY 1 KONSSTKSKRGK 14  
DB 34 KVEAKKTKKRGK 47  
RESULT 7  
HIT MOUSE STANDARD; PRT; 207 AA.  
ID -HIT MOUSE  
AC Q07133;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Histone H1t (Testicular H1 histone).  
GN H1FT OR H1T OR H1F3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=BALB/c; TISSUE=Leukocyte;  
RA MEDLINE=94060108; PubMed=8241275;  
RT Drabent B., Bode C., Doenecke D.;  
RT "Structure and expression of the mouse testicular H1 histone gene  
(H1t)."; Biochim. Biophys. Acta 1216:311-313(1993).  
RL [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=CD-1; TISSUE=Testis;  
RX van Wert J., Wright J., Wolfe S.A., Grimes S.R.;  
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: HISTONES H1 ARE NECESSARY FOR THE CONDENSATION OF  
CC NUCLEOSOME CHAINS INTO HIGHER ORDER STRUCTURES.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- DEVELOPMENTAL STAGE: THIS HISTONE IS A TESTIS-SPECIFIC H1 VARIANT  
CC THAT APPEARS DURING MEIOSIS IN SPERMATOGENESIS.  
CC -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L28753; AAB38417.1; -  
DR EMBL; X72805; CAA51325.1; -  
DR EMBL; U06232; AAA18359.1; -  
DR PIR; S33226; S33226.  
DR PIR; S43434; S43434.  
DR HSSP; P08287; 1GHC.  
DR MGD; MGI:107502; H1F3.  
DR InterPro; IPR001386; Histone\_H1/H5.  
DR InterPro; IPR00538; linker histone; 1.  
DR ProDom; PD000373; linkerhist\_N; 1.  
DR SMART; SM00526; H15; 1.  
DR KW Chromosomal protein; Nuclear protein; DNA-binding; Multigene family;  
KW Acetylation; Spermatogenesis; Testis.  
FT INIT MET 0 BY SIMILARITY.  
FT MOD RES 1 1 ACETYLATION (BY SIMILARITY).  
FT DOMAIN 37 110 GLOBULAR.  
FT CONFLICT 163 163 R -> G (IN REF. 1).  
SQ SEQUENCE 207 AA; 21508 MW; A3C847CA97976C44 CRC64;  
Query Match 54.1%; Score 40; DB 1; Length 207;  
Best Local Similarity 80.0%; Pred. No. 8.6;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
OY 6 KTSKRRGK 15  
DB 18 KPSSKRRGK 27  
RESULT 8  
VIF CAEVC STANDARD; PRT; 229 AA.  
ID -VIF CAEVC  
AC P33462;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE Virion infectivity factor (Q protein) (SOR protein).  
GN VIF.  
OS Caprine arthritis encephalitis virus (strain Cork) (CAEV).

OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11661;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91021037; PubMed=2171210;  
RA Saltarelli M., Querat G., Konings D.A.M., Vigne R., Clements J.E.;  
RT "Nucleotide sequence and transcriptional analysis of molecular clones  
of CAEV which generate infectious virus.",  
RL Virology 179:347-364(1990).  
CC -1- FUNCTION: DETERMINES VIRUS INFECTIVITY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M33677; AAA91827.1; -  
CC PIR; C45345; C45345.  
CC SEQUENCE 229 AA; 28051 MW; BF78AA3C4B23A101 CRC64;

Query Match 54.1%; Score 40; DB 1; Length 229;  
Best Local Similarity 57.1%; Pred. No. 9.5;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
Oy 2 QNSSKTSKRGRGK 15  
Db 2 QNSSRHQKKRNKK 15

RESULT 9  
YN05\_YEAST  
ID YN05\_YEAST STANDARD; PRT; 298 AA.  
AC P53843;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 34.5 kDa protein in PIK1-POL2 intergenic region.  
GN YNL265C OR N0809.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes;  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
STRAIN=S288C / FY1679;  
MEDLINE=96310631; PubMed=8740425;  
Sen-Gupta M., Lyck R., Fleig U., Niedenthal R.K., Hegemann J.H.;  
RT "The sequence of a 24,152 bp segment from the left arm of chromosome  
RT XIV from Saccharomyces cerevisiae between the BNI1 and the POL2  
RT genes."  
RL Yeast 12:505-514(1996).  
CC -1- SIMILARITY: TO HUMAN KIAA0174 AND S.POMBE SPCC285.02C.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X92494; -; NOT ANNOTATED\_CDS.  
DR EMBL; Z71542; CAA96172.1; -  
DR SGD; S0005209; YNL265C.  
DR InterPro; IPR005061; DUF292.  
DR Pfam; PF03398; DUF292; 1.  
KW Hypothetical protein.  
FT DOMAIN 190 196 POLY-SER.  
SQ SEQUENCE 298 AA; 34495 MW; F16A94BFC7716A1A CRC64;

Query Match 54.1%; Score 40; DB 1; Length 298;  
Best Local Similarity 53.3%; Pred. No. 12;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
Oy 1 KONSSTTSKRGRGK 15  
Db 264 KKGSEKTKTKRKX 278

RESULT 10  
YF16\_YEAST  
ID YF16\_YEAST STANDARD; PRT; 1233 AA.  
AC P43597;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Hypothetical 137.7 kDa protein in UGS1-FAB1 intergenic region.  
GN YFR016C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes;  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=95400292; PubMed=7670463;  
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,  
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,  
RA Yamazaki M., Tashiro H., Eki T.;  
RT "Analysis of the nucleotide sequence of chromosome VI from  
RT Saccharomyces cerevisiae."  
RL Nat. Genet. 10:261-268(1995).  
CC -1- SIMILARITY: SOME, TO MAMMALIAN NEUROFILAMENT TRIPLET M PROTEIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D50617; BAA09255.1; -  
CC DR SGD; S0001912; YFR016C.  
CC KW Hypothetical protein.  
SQ SEQUENCE 1233 AA; 137697 MW; C8A7CD2C6F0892F6 CRC64;

Query Match 54.1%; Score 40; DB 1; Length 1233;  
Best Local Similarity 50.0%; Pred. No. 53;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
Oy 1 KONSSTTSKRGRGK 14  
Db 186 KKNKKKKTNRGR 199

RESULT 11  
CYL1\_BOVIN  
ID CYL1\_BOVIN STANDARD; PRT; 667 AA.  
AC P35662;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Cytlicin I (Multiple-band polypeptide I).  
GN CYL1 OR CYL.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Testis;



```

RX MEDLINE=93359502; PubMed=8354692;
RA Hess H., Heid H., Franke W.W.;
RT "Molecular characterization of mammalian cyclin, a basic protein of
RL the sperm head cytoskeleton."
RL J. Cell Biol. 122:1043-1052(1993).
CC -1- FUNCTION: POSSIBLE ARCHITECTURAL ROLE DURING SPERMATOGENESIS. MAY
CC BE INVOLVED IN SPERMATID DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: CALYX; SPERM HEAD CYTOSKELETAL STRUCTURE.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- DEVELOPMENTAL STAGE: SPECIFIC TO LATE SPERMATOGENESIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z22779; CAA80456.1; -
CC PIR: S35913; S35913.
CC PIR: A40713; A40713.
KW Cytoskeleton; Structural protein; Repeat; Sperm; Spermatogenesis.
FT DOMAIN 287 569 9 APPROXIMATE TANDEM REPEATS.
FT REPEAT 287 305 1.
FT REPEAT 306 337 2.
FT REPEAT 338 368 3.
FT REPEAT 369 405 4.
FT REPEAT 406 442 5.
FT REPEAT 443 475 6.
FT REPEAT 476 516 7.
FT REPEAT 517 547 8.
FT REPEAT 548 569 9.
FT DOMAIN 617 667 PRO-RICH.
SQ SEQUENCE 667 AA; 74817 MW; CBF66EA462243D91 CRC64;

Query Match 52.7%; Score 39; DB 1; Length 667;
Best Local Similarity 53.3%; Pred. No. 42;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 KONSSTTSKRGKK 15
DB 596 KMPSRRITFKGKK 610

RESULT 12
UVRB_ECOLI STANDARD; PRT; 673 AA.
P07025;
01-APR-1988 (Rel. 07, Created)
01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Exci-nuclease ABC subunit B.
GN UVRB OR B0779 OR Z0998 OR ECS0857.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteri; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86176773; PubMed=3008099;
RA Backendorf C., Spalk H., Barbeiro A.P., van de Putte P.;
RT "Structure of the uvrB gene of Escherichia coli. Homology with other
RT DNA repair enzymes and characterization of the uvrB5 mutation.";
RL Nucleic Acids Res. 14:2877-2890(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86176755; PubMed=3515321;
RA Arlikan E., Kulkarni M.S., Thomas D.C., Sancar A.;
RT "Sequences of the E. coli uvrB gene and protein.";
RL Nucleic Acids Res. 14:2637-2650(1986).
RN [3]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Jim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
RN [6]
RP SEQUENCE OF 168-673 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K.,
RA Kashimoto K., Kimura S., Kitagawa M., Makino K., Masuda S.,
RA Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y.,
RA Nashimoto H., Nishio Y., Saito N., Sampei G., Seki Y., Tagami H.,
RA Takemoto K., Wada C., Yamamoto Y., Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:137-155(1996).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 619-673.
RX MEDLINE=20123894; PubMed=10631326;
RA Sohl M., Alexandrovich A., Moolenaar G., Visse R., Goosen N.,
RA Verneke X., Fontecilla-Camps J.C., Champness J., Sanderson M.R.;
RT "Crystal structure of Escherichia coli UvrB C-terminal domain, and a
RT model for UvrB-uvrC interaction."
RL FEBS Lett. 465:161-164(2000).
RN [8]
RP STRUCTURE BY NMR OF 619-673.
RX MEDLINE=99297571; PubMed=10371161;
RA Alexandrovich A., Sanderson M.R., Moolenaar G.F., Goosen N.,
RA Lane A.N.;
RT "NMR assignments and secondary structure of the UvrC binding domain of
RT UvrB."
RL FEBS Lett. 451:181-185(1999).
CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRB STIMULATES
CC THE ATPASE ACTIVITY OF UVRB IN THE PRESENCE OF UV-IRRADIATED
CC DOUBLE-STRANDED DNA. IT ALSO ENHANCES THE ABILITY OF UVRB TO BIND
CC TO UV-IRRADIATED DUPLEX DNA.
CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS, UVRB, UVRB AND UVRB.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE UVRB FAMILY.
CC -1- SIMILARITY: CONTAINS 1 UVR DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X03722; CAA27357.1; -  
DR EMBL; X03678; CAA27314.1; -  
DR EMBL; AE000180; AAC73866.1; -  
DR EMBL; AE005259; AAG55150.1; -  
DR EMBL; AP002553; BAB34280.1; -  
DR EMBL; D90716; BAA35437.1; -  
DR PIR; A23765; BVECUB.  
DR PDB; 1QOJ; 10-NOV-00.  
DR PDB; 1ESJ; 12-JUL-01.  
DR ECO2DBASE; C080.0; 6TH EDITION.  
DR Ecogen; EG1062; uvrB.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR001650; Helicase\_C.  
DR InterPro; IPR004807; uvrB.  
DR InterPro; IPR001943; uvrB/C.  
DR Pfam; PF00271; helicase\_C; 1.  
DR Pfam; PF02151; UVR; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR SMART; SM00490; HELIC; 1.  
DR TIGRFAM; TIGR00631; uvrB; 1.  
DR PROSITE; PS0151; UVR; 1.  
DR SOS response; Excision nuclease; DNA repair; ATP-binding;  
KW 3D-structure; Complete proteome.  
FT NP\_BIND 39 46 ATP (POTENTIAL).  
FT DOMAIN 633 668 UVR.  
FT SITE 630 631 CLEAVAGE.  
FT CONFLICT 477 477 H -> R (IN REF. 2).  
SQ SEQUENCE 673 AA; 76226 MW; 2F172045344FDAD7 CRC64;  
  
OY 2 QNSKTSKRRGK 14  
DB 601 QNTAKTKAKGRGK 613  
  
Query Match 52.7%; Score 39; DB 1; Length 673;  
Best Local Similarity 61.5%; Pred. No. 42;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
RESULT 13  
CENC\_MOUSE STANDARD; PRT; 906 AA.  
P49452;  
01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Centromere protein C (CENP-C) (Centromere autoantigen C).  
GN CENPC.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Swiss;  
RX MEDLINE=95048391; PubMed=7959789;  
RA McKay S., Thomson E., Cooke H.;  
RT "Sequence homologues and linkage group conservation of the human and  
RT mouse Cencp genes";  
RL Genomics 22:36-40(1994).  
CC -1- FUNCTION: COMPONENT OF THE INNER KINETOCHORE PLATE. REQUIRED FOR  
CC NORMAL KINETOCHORE ASSEMBLY.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U03113; AAC04314.1; -  
DR MGD; MGI:99700; Cenpc.  
KW Chromosomal protein; Nuclear protein; DNA-binding; Phosphorylation;  
KW Centromere.  
SQ SEQUENCE 906 AA; 102225 MW; 43D529B202E9D71E CRC64;  
  
OY 1 KONSSTTSKRRGK 15  
DB 451 KTNSTQTNKEKSGK 465  
  
Query Match 52.7%; Score 39; DB 1; Length 906;  
Best Local Similarity 46.7%; Pred. No. 57;  
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
  
RESULT 14  
IF41\_YEAST STANDARD; PRT; 952 AA.  
ID IF41\_YEAST  
AC P39935;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Eukaryotic initiation factor 4F subunit p150 (eIF-4F) (mRNA cap-  
DE binding protein complex subunit p150).  
GN TIF4631 OR YGRI62W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93330281; PubMed=8336723;  
RA Goyer C., Altman M., Lee H.S., Blanc A., Deshmukh M.,  
RA Woolford J.L., Trachsel H., Schenberg N.;  
RT "TIF4631 and TIF4632: two yeast genes encoding the  
RT high-molecular-weight subunits of the cap-binding protein complex  
RT (eukaryotic initiation factor 4F) contain an RNA recognition  
RT motif-like sequence and carry out an essential function.";  
RL Mol. Cell. Biol. 13:4860-4874(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c;  
RX MEDLINE=97435481; PubMed=9290212;  
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;  
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae  
RT chromosome VII.";  
RL Yeast 13:1077-1090(1997).  
CC -1- FUNCTION: INTERACTS WITH THE MRNA CAP STRUCTURE, M7GPPX. THIS  
CC INTERACTION IS REQUIRED FOR EFFICIENT RIBOSOME BINDING TO THE  
CC MRNA. TIF4631 IS PROBABLY ESSENTIAL WHEN TIF4632 IS MISSING.  
CC -1- SUBUNIT: THE CAP-BINDING PROTEIN COMPLEX IS COMPOSED OF AT LEAST  
CC TWO PROTEINS, A 24 Kda (TIF45) AND A 150-200 Kda SUBUNIT  
CC (TIF4631).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L16923; AAA02757.1; -  
DR EMBL; Z72947; CAA97184.1; -  
DR PIR; A48086; A48086.  
DR SGD; S0003394; TIF4631.  
DR InterPro; IPR003890; IF\_eIF4G.  
DR InterPro; IPR000504; RNA\_rec\_mot.

```

DR Pfam; PF02854; MIF4G; 1.
RA SMART; SM00543; MIF4G; 1.
KW Initiation factor; Protein biosynthesis; RNA-binding;
KW Multigene family.
FT DOMAIN 173 200 PRO/SER/THR-RICH.
FT DOMAIN 375 384 ALA/GLU-RICH.
FT DOMAIN 488 553 ARG/SER-RICH.
FT DOMAIN 873 899 ARG/SER-RICH.
FT DOMAIN 869 872 POLY-GLU.
FT CONFLICT 7 7 H -> Q (IN REF. 1).
FT CONFLICT 37 37 T -> N (IN REF. 1).
FT CONFLICT 110 110 Q -> K (IN REF. 1).
FT CONFLICT 207 207 R -> K (IN REF. 1).
FT CONFLICT 361 361 D -> E (IN REF. 1).
SQ SEQUENCE 952 AA; 107101 MW; 391256802F86118E CRC64;

Query Match 52.7%; Score 39; DB 1; Length 952;
Best Local Similarity 58.3%; Pred. No. 60;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
4 SSKTSKRRGKK 15
:::|||||:
Db 521 NSRTSSKRRSKR 532

```

```

RESULT 15
TOP2_SCHPO STANDARD; PRT; 1485 AA.
ID P08096; O74336;
AC 01-AUG-1988 (Rel. 08, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA topoisomerase II (EC 5.99.1.3).
GN TOP2 OR SPBC1A4.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87053875; PubMed=3023070;
RA Uemura T., Morikawa K., Yanagida M.;
RT "The nucleotide sequence of the fission yeast DNA topoisomerase II
RT gene: structural and functional relationships to other DNA
RT topoisomerases.";
RL EMBO J. 5:2355-2361(1986).
[2]
SEQUENCE FROM N.A.
STRAIN=972;
MEDLINE=21848401; PubMed=11859360;

```

```

RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voickaert G., Aert R., Robben J., Grymoprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

```

```

RA Cernutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [3]
RP DOMAINS.
RX MEDLINE=92049339; PubMed=1658625;
RA Shiozaki K., Yanagida M.;
RT "A functional 125-kDa core polypeptide of fission yeast DNA
RT topoisomerase II.";
RL Mol. Cell. Biol. 11:6093-6102(1991).
RN [4]
RP DOMAINS, AND PHOSPHORYLATION.
RX MEDLINE=93077665; PubMed=1332977;
RA Shiozaki K., Yanagida M.;
RT "Functional dissection of the phosphorylated termini of fission yeast
RT DNA topoisomerase II.";
RL J. Cell Biol. 119:1023-1036(1992).
CC -1- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC MAKES DOUBLE-STRAND BREAKS.
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: PHOSPHORYLATED AT MULTIPLE SITES AT BOTH EXTREMITIES OF THE
CC PROTEIN.
CC -1- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
CC RELAX ONLY NEGATIVE SUPERCOILS.
CC -1- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04326; CAA27857.1; ALT INIT.
DR EMBL; AL031174; CAA20107.1; -.
DR PIR; A24897; ISZPT2.
DR HSP; P06786; 1BGW.
DR InterPro; IPR003594; ATPbind ATPase.
DR InterPro; IPR003957; CBFA_NFYB_topis.
DR InterPro; IPR001241; DNA_topoisomII.
DR InterPro; IPR002205; DNA_topoisomIV.
DR Pfam; PF00204; DNA_gyraseB; 1.
DR Pfam; PF00521; DNA_topoisomIV; 1.
DR Pfam; PF00521; DNA_topoisomIV; 1.
DR Pfam; PF02518; HATPase c; 1.
DR Pfam; PF02518; HATPase c; 1.
DR PRINTS; PR00615; CCAATSUBUNTA.
DR PRINTS; PR00418; TP12FAMILY.
DR ProDom; PD000616; DNA_topoisomII; 1.
DR ProDom; PD000742; DNA_topoisomIV; 1.
DR SMART; SM00387; HATPase c; 1.
DR SMART; SM00433; TOP2c; 1.
DR SMART; SM00434; TOP4c; 1.
DR PROSITE; PS00177; TOPOISOMERASE II; 1.
DR PROSITE; PS00177; TOPOISOMERASE II; 1.
KW Isomerase; Topoisomerase; DNA-binding; ATP-binding; Phosphorylation;
KW Nuclear protein.
FT NP BIND 206 211 ATP (POTENTIAL).
FT ACT SITE 835 835 DNA CLEAVAGE (BY SIMILARITY).
FT CONFLICT 994 994 N -> I (IN REF. 1).
SQ SEQUENCE 1485 AA; 167891 MW; 6D88F76243361B2F CRC64;

```

```

Query Match 52.7%; Score 39; DB 1; Length 1485;
Best Local Similarity 53.3%; Pred. No. 94;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 KONSSTTSKRRGKK 15
:::|||||:
Db 27 KRKASTTSKSRRAK 41

```

Sun May 4 10:01:34 2003

us-09-757-982-5\_copy\_407\_421.rsp

Page 9

Search completed: May 1, 2003, 20:36:30  
Job time : 10.7531 secs

---





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:28:47 ; Search time 1.67364 Seconds

(without alignments)  
1846.697 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_407\_421

Perfect score: 74

Sequence: 1 KQNSSKTSKRGGK 15

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL 21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriophage:\*
- 17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description         |
|------------|-------|-------------|--------|-----------|---------------------|
| 1          | 74    | 100.0       | 455    | 4 Q9HCC4  | Q9HCC4 homo sapien  |
| 2          | 69    | 93.2        | 454    | 11 Q9ESL3 | Q9ESL3 mus musculu  |
| 3          | 48    | 64.9        | 241    | 11 Q9CRY1 | Q9CRY1 mus musculu  |
| 4          | 48    | 64.9        | 1117   | 11 Q88720 | Q88720 mus musculu  |
| 5          | 48    | 64.9        | 1122   | 11 Q8R388 | Q8R388 mus musculu  |
| 6          | 46    | 62.2        | 530    | 10 Q9XID7 | Q9XID7 arabidopsis  |
| 7          | 45    | 60.8        | 857    | 4 Q9BW51  | Q9BW51 homo sapien  |
| 8          | 44    | 59.5        | 425    | 5 Q24338  | Q24338 drosophila   |
| 9          | 43    | 58.1        | 124    | 10 Q9SGW6 | Q9SGW6 arabidopsis  |
| 10         | 43    | 58.1        | 270    | 13 Q9PUX7 | Q9PUX7 lampiris sp. |
| 11         | 43    | 58.1        | 556    | 10 Q8RYU7 | Q8RYU7 oryza sativ  |
| 12         | 43    | 58.1        | 683    | 5 Q09977  | Q09977 caenorhabdi  |
| 13         | 42    | 56.8        | 271    | 3 Q9UVP0  | Q9UVP0 candida gla  |
| 14         | 42    | 56.8        | 374    | 5 Q9XUG9  | Q9XUG9 caenorhabdi  |
| 15         | 42    | 56.8        | 1946   | 2 Q48545  | Q48545 lactobacill  |
| 16         | 41    | 55.4        | 425    | 5 Q26458  | Q26458 drosophila   |

|    |    |      |      |           |                     |
|----|----|------|------|-----------|---------------------|
| 17 | 41 | 55.4 | 453  | 10 Q9ZM89 | Q9ZM89 arabidopsis  |
| 18 | 41 | 55.4 | 761  | 4 Q9ULI9  | Q9ULI9 homo sapien  |
| 19 | 41 | 55.4 | 876  | 5 Q95ZS8  | Q95ZS8 caenorhabdi  |
| 20 | 41 | 55.4 | 925  | 5 Q95ZS7  | Q95ZS7 caenorhabdi  |
| 21 | 40 | 54.1 | 560  | 5 Q96139  | Q96139 plasmodium   |
| 22 | 40 | 54.1 | 3978 | 5 Q97236  | Q97236 plasmodium   |
| 23 | 39 | 52.7 | 82   | 17 Q8TJ10 | Q8TJ10 methanosarc  |
| 24 | 39 | 52.7 | 107  | 10 Q64841 | Q64841 arabidopsis  |
| 25 | 39 | 52.7 | 196  | 11 Q9D5N1 | Q9D5N1 mus musculu  |
| 26 | 39 | 52.7 | 241  | 10 Q9FEJ8 | Q9FEJ8 marchantia   |
| 27 | 39 | 52.7 | 254  | 10 Q8RXJ5 | Q8RXJ5 arabidopsis  |
| 28 | 39 | 52.7 | 273  | 13 Q9PUX0 | Q9PUX0 sarcocentro  |
| 29 | 39 | 52.7 | 312  | 3 Q06859  | Q06859 saccharomyc  |
| 30 | 39 | 52.7 | 312  | 3 Q06860  | Q06860 saccharomyc  |
| 31 | 39 | 52.7 | 319  | 16 Q8YF97 | Q8YF97 bruceella me |
| 32 | 39 | 52.7 | 325  | 10 Q9SNS9 | Q9SNS9 oryza sativ  |
| 33 | 39 | 52.7 | 328  | 5 Q9VBB3  | Q9VBB3 drosophila   |
| 34 | 39 | 52.7 | 328  | 5 Q9VBB3  | Q9VBB3 drosophila   |
| 35 | 39 | 52.7 | 382  | 12 Q72591 | Q72591 beet soil-b  |
| 36 | 39 | 52.7 | 401  | 4 Q9UFC3  | Q9UFC3 homo sapien  |
| 37 | 39 | 52.7 | 439  | 4 Q8TC11  | Q8TC11 homo sapien  |
| 38 | 39 | 52.7 | 441  | 12 Q9IH81 | Q9IH81 porcine epi  |
| 39 | 39 | 52.7 | 441  | 12 Q07499 | Q07499 porcine epi  |
| 40 | 39 | 52.7 | 446  | 10 Q94CT7 | Q94CT7 oryza sativ  |
| 41 | 39 | 52.7 | 485  | 5 Q9N821  | Q9N821 trypanosoma  |
| 42 | 39 | 52.7 | 546  | 5 Q22765  | Q22765 caenorhabdi  |
| 43 | 39 | 52.7 | 548  | 12 Q8UZC5 | Q8UZC5 chelonius in |
| 44 | 39 | 52.7 | 615  | 5 Q61609  | Q61609 acropora ml  |
| 45 | 39 | 52.7 | 627  | 4 Q9H8G3  | Q9H8G3 homo sapien  |

#### ALIGNMENTS

RESULT 1

Q9HCC4 PRELIMINARY; PRT; 455 AA.

AC Q9HCC4;

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE MLTK-beta (Similar to sterile-alpha motif and leucine zipper

DE containing kinase AZK) (Mixed lineage kinase) (Mixed lineage

DE kinase-related kinase MRK-beta).

GN MLTK.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Gotoh I., Adachi M., Nishida E.;

RT "Identification and Characterization of a Novel MAP Kinase Kinase

RT kinase, MLTK."

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=COLON;

RA Strausberg R.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Acton S.;

RT "MLK-mixed lineage kinase."

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=21950776; PubMed=11836244;

RA Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;

RT "MRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in

RT gamma-Radiation-induced Cell Cycle Arrest."

RL J. Biol. Chem. 277:13873-13882(2002).

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AB049734; BAB16445.1; -  
 DR EMBL; BC001401; AAH01401.1; -  
 DR EMBL; AF325454; AAK11615.1; -  
 DR EMBL; AF480462; AAL85892.1; -  
 DR HSSP; P12931; 1FMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TYRKc; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 455 AA; 51582 MW; E87DB84A4D58B752 CRC64;

Query Match 100.0%; Score 74; DB 4; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSCKTSKRRGKK 15  
 DB 407 KONSCKTSKRRGKK 421

RESULT 2  
 ID Q9ESL3 PRELIMINARY; PRT; 454 AA.  
 AC Q9ESL3;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE MLTK-beta.  
 GN ZAK OR MLTK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21264927; PubMed=11042189;  
 RA Gotch I., Adachi M., Nishida E.;  
 RT "Identification and Characterization of a Novel MAP Kinase Kinase Kinase, MLTK.";  
 JT Kinase, MLTK.";  
 J. Biol. Chem. 276:4276-4286(2001).  
 -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AB049732; BAB16443.1; -  
 DR HSSP; P12931; 1FMK.  
 DR MGD; MGI:1931274; Zak.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TYRKc; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 454 AA; 51366 MW; 35C2FC0D729D9395 CRC64;

Query Match .93.2%; Score 69; DB 11; Length 454;  
 Best Local Similarity .93.3%; Pred. No. 0.00052;  
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KONSCKTSKRRGKK 15

DB 407 KONSCKTSKRRGKK 421

RESULT 3  
 ID Q9CRY1 PRELIMINARY; PRT; 241 AA.  
 AC Q9CRY1;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE 13 days embryo head cDNA, RIKEN full-length enriched library,  
 clone:3100002020, full insert sequence (Fragment).  
 GN RBM6.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 Fleischmann W., Gaasterland T., Glissi C., King B., Kochiwa H.,  
 Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 Schriml L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 Guetincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Wittaker C., Wilming L.,  
 Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK013915; BAB29054.1; -  
 DR MGD; MGI:1338037; Rbm6.  
 FT NON TER 241 241  
 SQ SEQUENCE 241 AA; 27350 MW; 115150D1E98C6E71 CRC64;

Query Match 64.9%; Score 48; DB 11; Length 241;  
 Best Local Similarity 66.7%; Pred. No. 1.4;  
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KONSCKTSKRRGKK 15  
 DB 144 KONSCKTSKRRGKK 158

RESULT 4  
 ID O88720 PRELIMINARY; PRT; 1117 AA.  
 AC O88720;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE RNA binding protein.  
 GN RBM6 OR DEF-3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99280069; PubMed=10353602;  
 RA Drabkin H.A., West J., Hotfilder M., Heng Y.M., Erickson P., Calvo R.,

RA Dalmau J., Gemmili R.M., Sablitzky F.;  
RT "DEF-3, an RNA binding protein from the 3p21.3 homozygous deletion  
region in SCLC.";  
RL Oncogene 18:2589-2597(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99390180; PubMed=10460589;  
RA Hotfilder M., Sablitzky F.;  
RT "Def-2, -3, -4 and -8, novel mouse genes differentially expressed in  
the haemopoietic system.";  
RL Br. J. Haematol. 106:335-344(1999).  
DR EMBL; AJ006486; CAA07065.2; -  
DR MGD; MGI:1338037; Rbm6.  
DR InterPro; IPR000467; G\_patch.  
DR InterPro; IPR000504; RNA\_rec\_mot.  
DR InterPro; IPR003955; RRM\_2.  
DR InterPro; IPR008822; Znf\_C2H2.  
DR Pfam; PF01585; G\_patch; 1.  
DR Pfam; PF00076; xrm; 1.  
DR SMART; SM00443; G\_patch; 1.  
DR SMART; SM00360; RRM; 1.  
DR SMART; SM00362; RRM\_2; 1.  
DR SMART; SM00355; Znf\_C2H2; 1.  
DR PROSITE; PS50102; RRM; 2.  
DR PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 2.  
KW DNA-binding; Zinc-finger.  
SQ SEQUENCE 1117 AA; 128092 MW; F4485367E61AB107 CRC64;  
  
Query Match 64.9%; Score 48; DB 11; Length 1117;  
Best Local Similarity 66.7%; Pred. No. 6.2;  
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 KONSSTTSKRRGKK 15  
DB 837 KSNSSKETSRRDGE 851  
  
RESULT 5  
Q8R388 PRELIMINARY; PRT; 1122 AA.  
AC Q8R388;  
DT 01-JUN-2002 (Tremblrel. 21, Created)  
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Similar to RNA binding motif protein 6.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
NCBI\_TaxID=10090;  
[1]  
RP SEQUENCE FROM N.A.  
RA Strauberg R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC026129; AAH26129.1; -  
SQ SEQUENCE 1122 AA; 128836 MW; 8FE253C8E685CD5C CRC64;  
  
Query Match 64.9%; Score 48; DB 11; Length 1122;  
Best Local Similarity 66.7%; Pred. No. 6.2;  
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 KONSSTTSKRRGKK 15  
DB 837 KSNSSKETSRRDGE 851  
  
RESULT 6  
Q9XID7 PRELIMINARY; PRT; 530 AA.  
AC Q9XID7;  
DT 01-NOV-1999 (Tremblrel. 12, Created)  
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Hypothetical 59.7 kDa protein F23M19.1 (late embryogenesis abundant

DE Protein, putative).  
GN F23M19.1 OR F7P12.5.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,  
RA Lee J., Liu A., Li J., Kremenetskaia I., Lueros J., Gonzalez A.,  
RA Altafi H., Araujo R., Chao Q., Conn L., Conway A.B., Dunn P.,  
RA Hansen N., Huizar L., Kim C., Palm C., Rowley D., Shin P., Walker M.,  
RA Davis R.W., Ecker J.R., Federpsiel N.A., Theologis A.;  
RT "Arabidopsis thaliana chromosome 1 BAC F23M19 sequence.";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Theologis A.;  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
[3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RX MEDLINE=21016719; PubMed=11130712;  
RA Theologis A., Ecker J.R., Palm C.J., Federpsiel N.A., Kaul S.,  
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewart K.,  
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziali A.,  
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
RA Sakano H., Salzer S.L., Schwartz J.R., Shin P., Southwick A.M.,  
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
RA Uterback T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,  
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
thaliana.";  
RL Nature 408:816-820(2000).  
DR EMBL; AC007454; AAD39602.1; -  
DR EMBL; AC023913; AAG51902.1; -  
DR InterPro; IPR000379; Ser\_estrs\_site.  
KW Hypothetical protein.  
SQ SEQUENCE 530 AA; 59670 MW; EA07EC8C787A6589 CRC64;  
  
Query Match 62.2%; Score 46; DB 10; Length 530;  
Best Local Similarity 53.3%; Pred. No. 6.7;  
Matches 8; Conservative 6; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 KONSSTTSKRRGKK 15  
DB 460 EQDTNKTTLKQGRK 474  
  
RESULT 7  
Q9BW51 PRELIMINARY; PRT; 857 AA.  
AC Q9BW51;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE TTK protein kinase (DJ357D13.3) (PYT).  
GN TTK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;



```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 17-857 FROM N.A.
RA Tromans A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; BC000633; AAH00633.1; -.
DR EMBL; AL133475; CAB87580.1; -.
DR HSSP; Q00534; 1B18.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 857 AA; 97072 MW; 51F40A3CD1677AC5 CRC64;

Query Match 60.8%; Score 45; DB 4; Length 857;
Best Local Similarity 69.2%; Pred. No. 16;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 NSKTSKRRGKK 15
Db 845 SSKTFEKKRGKK 857

RESULT 8
ID Q24338 PRELIMINARY; PRT; 425 AA.
AC Q24338;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE ESC protein (SD03549p).
GN ESC OR CG14941.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_Taxid=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernick B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

```

```

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96016202; PubMed=7556071;
RA Gutjahr T., Frei E., Spicer C., Baumgartner S., White R.A., Noll M.;
RT "The Polcomb-group gene, extra sex combs, encodes a nuclear member of
the WD-40 repeat family.";
RL EMBO J. 14:4296-4306(1995).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazey R.G.,
RA Butenhof C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacle J.M., Park S., Pfeiffer B., Poon L., Punch E.,
RA Sequeira A., Sethi H., Snir E., Svirskas R.R., Twomey B., Wan K.H.,
RA Weinburg T., Zhang R., Zieran L.L., Rubin G.M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunco J., Pacle J.M., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL; AE003634; AAF53124.1; -.
DR EMBL; L41867; AA86427.1; -.
DR EMBL; AC006240; -. NOT ANNOTATED_CDS.
DR EMBL; AY069796; AAL39941.1; -.
DR TRANSFAC; T02156; -.
DR FlyBase; FBgn000588; esc.
DR InterPro; IPR000408; Reg_chr_condens.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR SMART; SM00320; WD40; 3.
DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
DR PROSITE; PS50082; WD_REPEATS_2; 2.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 425 AA; 47987 MW; 511C305E5DE86727 CRC64;

Query Match 59.5%; Score 44; DB 5; Length 425;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KNSKTSKRRGKK 15
Db 36 KSPSSSTRKRRGR 50

RESULT 9

```

Q9SGW6  
ID Q9SGW6 PRELIMINARY; PRT; 124 AA.  
AC Q9SGW6;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE FIN19.6.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chao Q., Brookes S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,  
RA Shin P., Tambunga G., Altafi H., Bei Q., Chin C., Chlou J., Choi E.,  
RA Conn L., Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B.,  
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,  
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
RA Thaverl A., Toriumi M., Vayberg M., Yu G., Federspiel N.A.,  
RA Theologis A., Ecker J.R.,  
RT "Genomic sequence for Arabidopsis thaliana BAC FIN19 from chromosome  
I.";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC009519; AAF19669.1;  
SQ SEQUENCE 124 AA; 14188 MW; 08DAB2EC13FE3D2A CRC64;  
QY Query Match 58.1%; Score 43; DB 10; Length 124;  
Best Local Similarity 50.0%; Pred. No. 5.6;  
Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;  
QY 2 QNSKTSKRRGK 15  
Db 97 ENATKTTTKKKK 110  
RESULT 10  
Q9PUX7  
ID Q9PUX7 PRELIMINARY; PRT; 270 AA.  
AC Q9PUX7;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
DE Mixed lineage leukemia-like protein (Fragment).  
GN MLL.  
OS Lampiris sp.  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Lampridiformes; Lamprididae; Lampiris.  
OX NCBI\_TaxID=94303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99398697; PubMed=10468597;  
RA Venkatesh B., Ning Y., Brenner S.;  
RT "Late changes in spliceosomal introns define clades in vertebrate  
evolution.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:10267-10271 (1999).  
DR EMBL; AF137231; AAD53454.1;  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 270 AA; 28613 MW; F7A86860A2C67F87 CRC64;  
QY Query Match 58.1%; Score 43; DB 13; Length 270;  
Best Local Similarity 61.5%; Pred. No. 12;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 3 NSSKTSKRRGK 15  
Db 177 NSHKNTGKRRGRE 189  
RESULT 11  
Q8RYU7

ID Q8RYU7 PRELIMINARY; PRT; 556 AA.  
AC Q8RYU7;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE B1033B05.8 protein.  
GN B1033B05.8.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC  
clone:B1033B05.";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP004223; BAB90658.1;  
SQ SEQUENCE 556 AA; 61490 MW; BAC85333021360E5 CRC64;  
QY Query Match 58.1%; Score 43; DB 10; Length 556;  
Best Local Similarity 50.0%; Pred. No. 24;  
Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KONSSTTSKRRGK 14  
Db 443 KDDAKTTNKXGK 456  
RESULT 12  
Q09977  
ID Q09977 PRELIMINARY; PRT; 683 AA.  
AC Q09977;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE Hypothetical 76.4 kDa protein.  
GN C17G10.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Johnson D.;  
RT "The sequence of C. elegans cosmid C17G10.";  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Waterston R.;  
RT "Direct Submission.";  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U28739; AAB93453.1;  
DR InterPro; IPR003582; SHKT.  
DR SMART; SM00254; SHKT; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 683 AA; 76395 MW; F0B7E15001672E8A CRC64;  
QY Query Match 58.1%; Score 43; DB 5; Length 683;  
Best Local Similarity 53.3%; Pred. No. 29;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
Q8RYU7

QY 1 KONSSTTSKRRGK 15  
|:|:|:|:|:|  
Db 557 KKKSKKNNKRRGK 571

## RESULT 13

Q9UVP0 PRELIMINARY; PRT; 271 AA.  
AC Q9UVP0;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE Hypothetical 30.7 kDa protein.  
GN YML041C.  
OS Candida glabrata (Yeast) (Torulopsis glabrata).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
OX NCBI\_TaxID=5478;  
[1]  
SEQUENCE FROM N.A.

STRAIN=CCN-1;  
Cheng S.J., Clancy C.J., Checkley M.A., Lewin A.S., Nguyen M.H.;  
"A Candida glabrata gene homologous to the gene encoding a  
hypothetical protein in the CAT2-AMD1 intergenic region of  
Saccharomyces cerevisiae.";  
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AF177905; AAD55397.1; -.  
DR Hypothetical protein.  
KW SEQUENCE 271 AA; 30728 MW; 0AE182D3679A0C9A CRC64;

Query Match 56.8%; Score 42; DB 3; Length 271;  
Best Local Similarity 53.3%; Pred. No. 18;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRGK 15  
|:|:|:|:|:|  
Db 39 KKSNNKITSSRNAK 53

## RESULT 14

Q9XUG9 PRELIMINARY; PRT; 374 AA.

AC Q9XUG9;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE C43F9.6 protein.  
GN C43F9.6.

OC Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
[1]

SEQUENCE FROM N.A.  
RP Mortimore B.J.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases;  
[2]  
SEQUENCE FROM N.A.  
RP MEDLINE=99069613; PubMed=9851916;  
RX none;

RA "Genome sequence of the nematode C.elegans: A platform for  
RT investigating biology.";  
RT Science 282:2012-2018 (1998).  
RL EMBL; Z82262; CAB05149.1; -.  
DR InterPro; IPR000402; Na/K\_ATPase\_beta.  
DR Pfam; PF00287; Na\_K-ATPase; 1.  
DR SEQUENCE 374 AA; 41940 MW; D751E6610D8EBE72 CRC64;

Query Match 56.8%; Score 42; DB 5; Length 374;  
Best Local Similarity 69.2%; Pred. No. 24;  
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRG 13  
|:|:|:|:|:|  
Db 20 KONSSTTSKRRVG 32

## RESULT 15

Q48545 PRELIMINARY; PRT; 1946 AA.  
AC Q48545;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Proteinase precursor.  
GN PRTB.  
OS Lactobacillus delbrueckii.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Lactobacillaceae; Lactobacillus.  
OX NCBI\_TaxID=1584;  
[1]

SEQUENCE FROM N.A.  
RP MEDLINE=96236017; PubMed=8655480;  
RX Gilbert C., Atlan D., Blanc B., Portailier R., Germond J.E.,  
RA Lapiere L., Mollet B.;  
RT "A new cell surface proteinase: sequencing and analysis of the prtb  
gene from Lactobacillus delbrueckii subsp. bulgaricus.";  
RL J. Bacteriol. 178:3059-3065 (1996).

[2]  
SEQUENCE FROM N.A.  
RP Germond J.E.;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; L48487; AAC41529.1; -.  
DR HSSP; P00782; 1SUE.  
DR InterPro; IPR003137; PA.  
DR InterPro; IPR000209; Peptidase\_S8.  
DR Pfam; PF02225; PA; 1.  
DR Pfam; PF00082; Peptidase\_S8; 3.  
DR PRINTS; PR00723; SUBTILISIN.  
DR PROSITE; PS00136; SUBTILASE\_ASP; UNKNOWN\_1.  
DR PROSITE; PS00138; SUBTILASE\_SER; 1.  
KW Signal.  
FT SIGNAL 1 34 POTENTIAL.  
FT CHAIN 193 1946 PROTEINASE.  
FT SEQUENCE 1946 AA; 212315 MW; 21EF17D02E79C6A0 CRC64;

Query Match 56.8%; Score 42; DB 2; Length 1946;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRGK 15  
|:|:|:|:|:|  
Db 1789 KKTDSKTSKSSAKK 1803

Search completed: May 1, 2003, 20:37:51  
Job time : 4.67364 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:12:12 ; Search time 2.88703 Seconds

(without alignments)  
692.324 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_407\_421

Perfect score: 74

Sequence: 1 KONSSTTSKRGGK 15

Scoring table: BLOSUM62

Searched: 908470 seqs, 133250620 residues

al number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID       | Description            |
|------------|-------|-------------|--------|----|----------|------------------------|
| 1          | 74    | 100.0       | 124    | 22 | AAB65552 | C-terminus specific    |
| 2          | 74    | 100.0       | 455    | 21 | AAB18657 | A human regulator      |
| 3          | 74    | 100.0       | 455    | 21 | AAV83278 | Human survival reg     |
| 4          | 74    | 100.0       | 455    | 21 | AAV84321 | A human cardiovascular |
| 5          | 74    | 100.0       | 473    | 22 | AAV25322 | Human survival reg     |
| 6          | 64    | 86.5        | 14     | 21 | AAV83280 | Human survival reg     |
| 7          | 44    | 59.5        | 425    | 22 | ABB69961 | Drosophila melanog     |
| 8          | 43    | 58.1        | 51     | 21 | AAG28664 | Arabidopsis thalia     |
| 9          | 43    | 58.1        | 124    | 21 | AAG41571 | Arabidopsis thalia     |
| 10         | 42    | 56.8        | 179    | 22 | ABG14321 | Novel human diagno     |

|    |    |      |      |    |          |                    |
|----|----|------|------|----|----------|--------------------|
| 11 | 42 | 56.8 | 179  | 22 | ABG17137 | Novel human diagno |
| 12 | 42 | 56.8 | 1946 | 19 | AAW47273 | Lactobacillus bulg |
| 13 | 41 | 55.4 | 18   | 21 | AAV65802 | Kinase TRK mutant  |
| 14 | 41 | 55.4 | 425  | 22 | AAU29308 | Human PRO polypept |
| 15 | 41 | 55.4 | 734  | 22 | AAV78679 | Human protein seq  |
| 16 | 41 | 55.4 | 1113 | 22 | ABG03554 | Novel human diagno |
| 17 | 41 | 55.4 | 1998 | 22 | ABG19486 | Novel human diagno |
| 18 | 40 | 54.1 | 78   | 23 | ABP00883 | Human ORFX protein |
| 19 | 40 | 54.1 | 90   | 22 | AAV76540 | Human immune/haema |
| 20 | 40 | 54.1 | 104  | 20 | AAV76540 | Human ovarian tumo |
| 21 | 40 | 54.1 | 210  | 22 | AAV74869 | Human colon cancer |
| 22 | 40 | 54.1 | 320  | 22 | ABG21345 | Novel human diagno |
| 23 | 40 | 54.1 | 560  | 21 | AAV18178 | Plasmodium falcipa |
| 24 | 39 | 52.7 | 106  | 21 | AAV57182 | Human prostate can |
| 25 | 39 | 52.7 | 113  | 22 | ABG25930 | Novel human diagno |
| 26 | 39 | 52.7 | 184  | 22 | AAV75010 | Human breast cance |
| 27 | 39 | 52.7 | 200  | 22 | AAV63344 | Human breast cance |
| 28 | 39 | 52.7 | 242  | 22 | AAV63343 | Human breast cance |
| 29 | 39 | 52.7 | 323  | 22 | AAV39441 | Human polypeptide  |
| 30 | 39 | 52.7 | 328  | 22 | ABB61583 | Drosophila melanog |
| 31 | 39 | 52.7 | 328  | 22 | ABB67466 | Drosophila melanog |
| 32 | 39 | 52.7 | 395  | 22 | AAV63337 | Human breast cance |
| 33 | 39 | 52.7 | 401  | 22 | AAV32047 | Novel human secret |
| 34 | 39 | 52.7 | 524  | 22 | AAV15947 | Human novel secret |
| 35 | 39 | 52.7 | 524  | 22 | AAV16394 | Human novel secret |
| 36 | 39 | 52.7 | 627  | 22 | AAV94704 | Human protein sequ |
| 37 | 39 | 52.7 | 673  | 21 | AAV58042 | Escherichia coli U |
| 38 | 39 | 52.7 | 1035 | 22 | ABG05319 | Novel human diagno |
| 39 | 39 | 52.7 | 1592 | 22 | AAV38645 | Human polypeptide  |
| 40 | 39 | 52.7 | 1625 | 22 | AAV40431 | Human polypeptide  |
| 41 | 39 | 52.7 | 1898 | 23 | ABB07352 | Amino acid sequenc |
| 42 | 38 | 51.4 | 137  | 22 | ABG02379 | Novel human diagno |
| 43 | 38 | 51.4 | 148  | 21 | AAV08005 | Arabidopsis thalia |
| 44 | 38 | 51.4 | 148  | 21 | AAV45485 | Arabidopsis thalia |
| 45 | 38 | 51.4 | 159  | 21 | AAV08004 | Arabidopsis thalia |

ALIGNMENTS

|          |             |   |
|----------|-------------|---|
| RESULT 1 | AAV65552    | standard; Protein; 124 AA.  |
| ID       | AAV65552;   |   |
| AC       | 27-MAR-2001 | (first entry)   |
| DT       |             |   |
| XX       |             |   |
| DE       |             | C-terminus specific to novel human protein kinase MLK4B.                |
| XX       |             |   |
| KW       |             | Human; protein kinase; antirheumatic; antisclerotic; immunosuppressive; |
| KW       |             | cardiant; renal; antiinflammatory; antiaesthetic; osteopathic;          |
| KW       |             | dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;   |
| KW       |             | immune disorder; cardiovascular disease; neurodegenerative disease;     |
| KW       |             | cancer; autoimmune disorder; stroke; inflammatory bowel disease;        |
| KW       |             | inflammatory pelvic disease; multiple sclerosis; psoriasis.             |
| OS       |             | Homo sapiens.   |
| XX       |             |   |
| PN       |             | WO200073469-A2.   |
| XX       |             |   |
| PD       |             | 07-DEC-2000.  |
| XX       |             |   |
| PF       |             | 26-MAY-2000; 2000WO-US14842.  |
| XX       |             |   |
| PR       |             | 28-MAY-1999; 99US-0136503.  |
| XX       |             |   |
| PA       |             | (SUGEN-) SUGEN INC.   |
| XX       |             |   |
| PI       |             | Plowman GD, Martinez R, Whyte D, Sudersanam S;                          |
| XX       |             |   |
| DR       |             | WPI; 2001-032161/04.  |
| XX       |             |   |





CC inflammatory disorder such as Addison's disease, acquired  
CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,  
CC rheumatoid arthritis, microbial infection and trauma.  
XX  
SQ Sequence 455 AA;

Query Match 100.0%; Score 74; DB 21; Length 455;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
ID AAY83278  
DB 407 KONSSTTSKRRGKK 421

RESULT 3  
AAY83278  
ID AAY83278 standard; Protein; 455 AA.  
XX AAY83278;

DT 16-AUG-2000 (first entry)  
XX  
DE Human survival regulating kinase (SRK).  
XX

KW Survival regulating kinase; SRK; Raf; MBP; BAD; MAPKK; MAPKK;  
KW HAX-1; protein kinase; autophosphorylation; cell growth; regulation;  
KW apoptosis; cell survival; nuclear targeting; tumour; human;  
KW autoimmune disease.  
XX

OS Homo sapiens.  
XX

PN WO200022142-A2.  
XX

PD 20-APR-2000.  
XX

PF 20-SEP-1999; 99WO-US22008.  
XX

PR 13-OCT-1998; 98US-0104088.  
XX

PA (ONYX-) ONYX PHARM INC.  
XX

PI Rugstjeri R, Callow M, Diaz P;  
XX

DR WPI; 2000-317994/27.  
XX

DR N-PSDB; AAZ93783.  
XX

PT Novel human survival regulating kinase polypeptide for screening agents  
XX which modulate biological pathways associated with SRK useful in  
XX treating autoimmune diseases, tumors and apoptosis-related disorders  
XX  
PS Claim 4; Figure 2; 62pp; English.  
XX

CC Survival regulating kinases (SRK) are a class of proteins involved in  
CC cell signal transduction pathways such as mitogen-activated protein  
CC kinase pathways. A protein kinase activity means that the SRK can  
CC catalyze a reaction in which a phosphate group is transferred from a  
CC phosphate donor to a hydroxyl side chain of a serine or threonine.  
CC Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase  
CC activity is similar to that of a MAPKK such as Raf. has a range of  
CC other activities including a cell growth-regulatory activity, a cell  
CC survival promoting activity, a HAX-1 binding activity, an apoptosis  
CC suppressing activity a MAPKK activation or stimulatory activity, a cell  
CC nuclear targeting activity and a SRK-specific immunogenic activity.  
CC SRK is useful for identifying agents which modulate cellular  
CC transformations mediated by Ras and SRK and agents that modulate the  
CC apoptosis suppression activity of SRK. This information may be useful  
CC in the treatment of autoimmune diseases, tumours and apoptosis  
CC related disorders.  
XX  
SQ Sequence 455 AA;

Query Match 100.0%; Score 74; DB 21; Length 455;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
DB 407 KONSSTTSKRRGKK 421

RESULT 4  
AAY84321  
ID AAY84321 standard; Protein; 455 AA.  
XX  
AC AAY84321;  
XX

DT 12-JUL-2000 (first entry)  
XX

DE A human cardiovascular system associated protein kinase-2.  
XX

KW Human; cardiovascular system associated protein kinase-2; CSAPK-2;  
KW signalling pathway; cell growth; cell differentiation; gene mapping;  
KW tissue typing; forensic identification; cardiovascular disease;  
KW congestive heart failure; transgenic animal.  
XX

OS Homo sapiens.  
XX

PN WO200014212-A1.  
XX

PD 16-MAR-2000.  
XX

PF 09-SEP-1999; 99WO-US20631.  
XX

PR 09-SEP-1998; 98US-0099657.  
XX

PR 29-SEP-1998; 98US-0163115.  
XX

PA (MILL-) MILLENNIUM PHARM INC.  
XX

PI Acton S;  
XX

DR WPI; 2000-271053/23.  
XX

DR N-PSDB; AAZ99726, AAZ99727.  
XX

PT New nucleic acid encoding cardiovascular system associated protein  
XX kinase, used e.g. for diagnosis, treatment and prevention of  
XX cardiovascular disease  
XX

PS Claim 2; Fig 2; 163pp; English.  
XX

CC The present sequence represents a human cardiovascular system associated  
CC protein kinase-2 (CSAPK-2). CSAPK polypeptides are involved in signalling  
CC pathways associated with cell growth and differentiation. The CSAPK  
CC polypeptides and polynucleotides are used to screen for agents that  
CC specifically modulate CSAPK, which are potential therapeutic agents.  
CC They are also used for diagnosis, prognosis or monitoring of  
CC CSAPK-related diseases, gene mapping, tissue typing and forensic  
CC identification, and for treating or preventing disorders associated  
CC with aberrant CSAPK expression or activity, especially cardiovascular  
CC diseases such as congestive heart failure. They can also be used in  
CC pharmacogenomics. The CSAPK polynucleotide may also be used to generate  
CC transgenic animals.  
XX  
SQ Sequence 455 AA;

Query Match 100.0%; Score 74; DB 21; Length 455;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
DB 407 KONSSTTSKRRGKK 421

RESULT 5

AAM25322  
ID AAM25322 standard; Protein; 473 AA.  
XX  
AC AAM25322;  
XX  
DT 16-OCT-2001 (first entry)  
XX  
DE Human protein sequence SEQ ID NO:837.  
XX  
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
KW antibacterial; endocrine; cardiant; central nervous system; virucide;  
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
KW antiaggregant; haemostatic; vulnery; antilucer; osteopathic; eczema;  
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;  
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200153455-A2.  
XX  
PD 26-JUL-2001.  
XX  
PF 22-DEC-2000; 2000WO-US35017.  
XX  
PR 23-DEC-1999; 99US-0471275.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
XX  
PR 25-APR-2000; 2000US-0552317.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-457603/49.  
XX  
DR N-PSDB; AAM99263.  
XX  
XX  
PT Isolated human polynucleotides encoding polypeptides, useful for the  
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -  
XX  
XX  
PS Claim 20; Page 191; 1217bp; English.  
AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
AAM25963. The proteins can have activities based on the tissues and  
cells they are expressed in, such as: antiinflammatory; antirheumatic;  
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;  
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;  
CC antilucer; osteopathic; dermatological; antiallergic; antiasthmatic;  
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;  
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
CC encoding them can be used in gene therapy, antisense therapy and vaccine  
CC production. The proteins and polynucleotides are useful for screening for  
CC agonists or antagonists of a protein and for the treatment and diagnosis  
CC of disorders associated with the activity of a protein e.g. inflammation,  
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders.  
XX  
SQ Sequence 473 AA;

Query Match 100.0%; Score 74; DB 22; Length 473;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KONSSTTSKRRGK 15  
DB 425 KONSSTTSKRRGK 439  
RESULT 6  
ID AAY83280  
XX  
AC AAY83280;  
XX  
DT 16-AUG-2000 (first entry)  
XX  
DE Human survival regulating kinase (SRK) antigenic peptide.  
XX  
KW Survival regulating kinase; SRK; Raf; MBP; BAD; MAPKK; MAPKK;  
KW HAX-1; protein kinase; autophosphorylation; cell growth; regulation;  
KW apoptosis; cell survival; nuclear targeting; tumour; human;  
KW autoimmune disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200022142-A2.  
XX  
PD 20-APR-2000.  
XX  
PF 20-SEP-1999; 99WO-US22008.  
XX  
PR 13-OCT-1998; 98US-0104088.  
XX  
PA (ONYX-) ONYX PHARM INC.  
XX  
PI Ruggieri R, Callow M, Diaz P;  
XX  
DR WPI; 2000-317994/27.  
XX  
PT Novel human survival regulating kinase polypeptide for screening agents  
PT which modulate biological pathways associated with SRK useful in  
PT treating autoimmune diseases, tumors and apoptosis-related disorders  
XX  
XX  
PS Claim 53; Page 6; 62pp; English.  
XX  
CC Survival regulating kinases (SRK) are a class of proteins involved in  
CC cell signal transduction pathways such as mitogen-activated protein  
CC kinase pathways. A protein kinase activity means that the SRK can  
CC catalyse a reaction in which a phosphate group is transferred from a  
CC phosphate donor to a phosphate acceptor amino acid residue,  
CC preferably the hydroxyl side chain of a serine or threonine.  
CC Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase  
CC activity is similar to that of a MAPKK such as Raf. has a range of  
CC other activities including a cell growth-regulatory activity, a cell  
CC survival promoting activity, a HAX-1 binding activity, an apoptosis  
CC suppressing activity a MAPKK activation or stimulatory activity, a  
CC nuclear targeting activity and a SRK-specific immunogenic activity.  
CC SRK is useful for identifying agents which modulate cellular  
CC transformations mediated by Ras and SRK and agents that modulate the  
CC apoptosis suppression activity of SRK. This information may be useful  
CC in the treatment of autoimmune diseases, tumours and apoptosis  
CC related disorders.  
XX  
SQ Sequence 14 AA;  
Query Match 86.5%; Score 64; DB 21; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KONSSTTSKRRG 13  
DB 2 KONSSTTSKRRG 14

RESULT 7  
ID ABB69961 standard; Protein; 425 AA.  
XX  
AC ABB69961;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 36675.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PF 27-SEP-2001.  
XX  
PR 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
XX 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
XX N-PSDB; ABL14064.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Disclosure; SEQ ID NO 36675; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 425 AA;  
XX  
Query Match 59.5%; Score 44; DB 22; Length 425;  
Best Local Similarity 60.0%; Pred. No. 21;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
OY 1 KONSSTTSKRGKK 15  
Db 36 KSPSSSTRSKRGRR 50  
RESULT 8  
ID AAG28664 standard; Protein; 51 AA.  
XX  
AC AAG28664;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33967.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134376.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.



PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 07-OCT-1999; 99US-0157865.  
PR 08-OCT-1999; 99US-0158029.  
PR 12-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 58.1%; Score 43; DB 21; Length 51;  
Best Local Similarity 50.0%; Pred. No. 3.3;  
Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 2 ONSKTSKRRGKK 15  
Db 24 ENATKTTTKKKKK 37

RESULT 9  
AAG41571  
ID AAG41571 standard; Protein; 124 AA.

XX AC AAG41571;  
XX DT 18-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 51740.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.

XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX XX

PD 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
PF  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.

|    |              |               |
|----|--------------|---------------|
| PR | 23-SEP-1999; | 99US-0155486. |
| PR | 24-SEP-1999; | 99US-0155659. |
| PR | 28-SEP-1999; | 99US-0156458. |
| PR | 29-SEP-1999; | 99US-0156596. |
| PR | 04-OCT-1999; | 99US-0157117. |
| PR | 05-OCT-1999; | 99US-0157753. |
| PR | 06-OCT-1999; | 99US-0157865. |
| PR | 07-OCT-1999; | 99US-0158029. |
| PR | 08-OCT-1999; | 99US-0158232. |
| PR | 12-OCT-1999; | 99US-0158369. |
| PR | 13-OCT-1999; | 99US-0159293. |
| PR | 13-OCT-1999; | 99US-0159294. |
| PR | 13-OCT-1999; | 99US-0159295. |
| PR | 14-OCT-1999; | 99US-0159329. |
| PR | 14-OCT-1999; | 99US-0159330. |
| PR | 14-OCT-1999; | 99US-0159331. |
| PR | 14-OCT-1999; | 99US-0159637. |
| PR | 14-OCT-1999; | 99US-0159638. |
| PR | 18-OCT-1999; | 99US-0159584. |
| PR | 21-OCT-1999; | 99US-0160741. |
| PR | 21-OCT-1999; | 99US-0160767. |
| PR | 21-OCT-1999; | 99US-0160768. |
| PR | 21-OCT-1999; | 99US-0160770. |
| PR | 21-OCT-1999; | 99US-0160814. |
| PR | 21-OCT-1999; | 99US-0160815. |
| PR | 22-OCT-1999; | 99US-0160980. |
| PR | 22-OCT-1999; | 99US-0160981. |
| PR | 22-OCT-1999; | 99US-0160989. |
| PR | 25-OCT-1999; | 99US-0161404. |
| PR | 25-OCT-1999; | 99US-0161405. |
| PR | 25-OCT-1999; | 99US-0161406. |
| PR | 26-OCT-1999; | 99US-0161359. |
| PR | 26-OCT-1999; | 99US-0161360. |
| PR | 26-OCT-1999; | 99US-0161361. |
| PR | 28-OCT-1999; | 99US-0161920. |
| PR | 28-OCT-1999; | 99US-0161992. |
| PR | 28-OCT-1999; | 99US-0161993. |
| PR | 29-OCT-1999; | 99US-0162142. |

|                       |        |                |        |               |
|-----------------------|--------|----------------|--------|---------------|
| Query Match           | 58.1%; | Score 43;      | DB 21; | Length 124;   |
| Best Local Similarity | 50.0%; | Pred. No. 8.6; |        |               |
| Matches               | 7;     | Conservative   | 6;     | Mismatches 1; |
|                       |        |                |        | Indels 0;     |
|                       |        |                |        | Gaps 0;       |

```
QY      2 QNSSKTTSKRRGKK 15
          ::::|::|::|
Db      97 ENATKTTTKKKKKK 110
```

RESULT 10  
ABG14321 standard; Protein; 179 AA.

AC ABG14321;

DT 18-FEB-2002 (first entry)

Novel human diagnostic protein #14312.

**KW** Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN W0200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;  
PI  
XX WPI; 2001-639362/73.  
DR N-PSDB; AAS78508.  
DR

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

Claim 20; SEQ ID No 44680; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](ftp.wipo.int/pub/published_pct_sequences).

SQ Sequence 179 AA;

|                       |        |               |        |                                 |
|-----------------------|--------|---------------|--------|---------------------------------|
| Query Match           | 56.8%; | Score 42;     | DB 22; | Length 179;                     |
| Best Local Similarity | 57.1%; | Pred. No. 19; |        |                                 |
| Matches               | 8;     | Conservative  | 2;     | Mismatches 4; Indels 0; Gaps 0; |

```

Oy      2  QNSSKTTSKRRGKK 15
          ||||:|  : |||
Db      11  QNSSETIKNNKGKK 24

```

## RESULT 11

ABG17137 standard; Protein; 179 AA.

AC ABG17137;

DT 18-FEB-2002 (first entry)

Novel human diagnostic protein #17128.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic; genetic disorders; diagnostic; genetic disorder

XX  
OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

XX  
PA (HYSE-) HYSEQ INC.  
...

PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.  
DR N-PSDB; AAS81324.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity

PS Claim 20; SEQ ID No 47496; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations in  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG0010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 179 AA;

Query Match 56.8%; Score 42; DB 22; Length 179;  
Best Local Similarity 57.1%; Pred. No. 19;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ONSKTTSKRRGK 15  
DB 11 ONSSETIKKNGK 24

RESULT 12  
AAW47273

AAW47273 standard; Protein; 1946 AA.

AAW47273;

22-MAY-1998 (first entry)

Lactobacillus bulgaricus protease P.

Protease P; transformed microorganism; recombinant production;  
PrtP; yoghurt; acidulated milk; cheese.

Lactobacillus bulgaricus.

EP810289-A1.

03-DEC-1997.

29-MAY-1996; 96EP-0201495.

29-MAY-1996; 96EP-0201495.

(NEST ) SOC PROD NESTLE SA.

Germond JE, Lapierre L, Mollet B;

WPI; 1998-011060/02.  
N-PSDB; AAV15586.

XX Recombinant Lactobacillus bulgaricus PrtP protease - and DNA for  
PT transforming microorganisms for making fermented dairy products  
XX Claim 1; Pages 9-17; 28pp; English.

XX The present sequence is Lactobacillus bulgaricus protease P  
CC (PrtP). A microorganism transformed with the PrtP DNA can be used  
CC for the recombinant production of the enzyme, or in the manufacture  
CC of fermented dairy products, e.g. yoghurt, acidulated milk and  
CC cheese, especially where: (a) the microorganism expresses a  
CC recombinant PrtP derivative that is temperature or pH sensitive,  
CC preferably having no more than 80% of wild type protease activity  
CC under storage conditions and at least 90% of wild type protease  
CC activity under fermentation conditions; or (b) the microorganism  
CC contains PrtP DNA under the control of a temperature or pH  
CC sensitive promoter.

XX Sequence 1946 AA;

Query Match 56.8%; Score 42; DB 19; Length 1946;  
Best Local Similarity 60.0%; Pred. No. 2.4e+02;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KONSKTTSKRRGK 15  
DB 1789 KRTDSKTTSKSAKK 1803

RESULT 13  
AAY65802

AAY65802 standard; Peptide; 18 AA.

AAY65802;

10-FEB-2000 (first entry)

Kinase TTK mutant peptide 11.

Human; frameshift mutant; T cell response; tumour; treatment; cancer;

mutain.

Homo sapiens.  
OS Synthetic.

WO9958552-A2.

18-NOV-1999.

03-MAY-1999; 99WO-NO00143.

08-MAY-1998; 98NO-0002097.

(NHYD ) NORSK HYDRO AS.

Gaudernack G, Eriksen JA, Moller M, Gjertsen MK, Saeterdal I;

WPI; 2000-039064/03.

XX New peptides derived from genes with frameshift mutations, used to  
PT develop products for the treatment and prophylaxis of cancers -  
XX Claim 13; Page 27; 166pp; English.

XX Peptides AAY65684-Y66142 are fragments of mutant proteins arising from a  
CC frameshift mutation in a gene from a cancer cell. The peptides are  
CC characterised in that they:

CC (i) are at least 8 amino acids long and a fragment of a mutant protein  
CC arising from a frameshift mutation in a gene of a cancer cell;  
CC (ii) consist of at least one amino acid of the mutant part of a protein  
CC sequence encoded by the gene;  
CC (iii) comprise 0-10 amino acid from the carboxyl terminus of the normal  
CC part of the protein sequence preceding the amino terminus of the mutant





XX OS Homo sapiens.  
XX PN WO200157190-A2.  
XX PD 09-AUG-2001.  
XX PF 05-FEB-2001; 2001WO-US04098.  
XX PR 03-FEB-2000; 2000US-0496914.  
PR 27-APR-2000; 2000US-0560875.  
PR 20-JUN-2000; 2000US-0598075.  
PR 19-JUL-2000; 2000US-0620325.  
PR 01-SEP-2000; 2000US-0654936.  
PR 15-SEP-2000; 2000US-0663561.  
PR 20-OCT-2000; 2000US-0693325.  
PR 30-NOV-2000; 2000US-0728422.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX  
XX WPI; 2001-476283/51.  
XX N-PSDB; AAK51812.  
XX  
XX Nucleic acids encoding polypeptides with cytokine-like activities,  
XX useful in diagnosis and gene therapy -  
XX  
XX Claim 20; Page 3585-3586; 6221pp; English.  
XX  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
XX cytokine, cell proliferation or cell differentiation or which may induce  
XX production of other cytokines in other cell populations. The  
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or  
XX peptide therapy. The polypeptides have various cytokine-like activities,  
XX e.g. stem cell growth factor activity, haematopoiesis regulating  
XX activity, tissue growth factor activity, immunomodulatory activity and  
XX activin/inhibin activity and may be useful in the diagnosis and/or  
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and  
XX inflammation.  
XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
XX (AAM80020) are omitted as the relevant pages from the sequence listing  
XX were missing at the time of publication.  
XX  
XX Sequence 734 AA;  
XX  
XX Query Match 55.4%; Score 41; DB 22; Length 734;  
XX Best Local Similarity 53.3%; Pred.No. 1.3e+02;  
XX Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
XX  
XX 1 KONSSTTSKRRGKK 15  
XX | : | | | | : | : | :  
XX Db 552 KEKASKTTTVRRKR 566

Search completed: May 1, 2003, 20:35:38  
Job time : 4.88703 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: May 1, 2003, 20:38:08 ; Search time 0.962343 Seconds  
(without alignments)  
1344.947 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_407\_421

Perfect score: 74  
Sequence: 1 KQNSKTSKRGGK 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

Optimal number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID                | Description        |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1          | 74    | 100.0       | 455    | 10 US-09-757-982-5   | Sequence 5, Appli  |
| 2          | 45    | 60.8        | 841    | 9 US-10-081-119-14   | Sequence 14, Appli |
| 3          | 42    | 56.8        | 312    | 10 US-09-801-368-260 | Sequence 260, App  |
| 4          | 41    | 55.4        | 425    | 9 US-10-174-590-570  | Sequence 570, App  |
| 5          | 41    | 55.4        | 425    | 9 US-10-176-758-570  | Sequence 570, App  |
| 6          | 41    | 55.4        | 425    | 9 US-10-175-737-570  | Sequence 570, App  |
| 7          | 41    | 55.4        | 425    | 9 US-10-173-706-570  | Sequence 570, App  |
| 8          | 41    | 55.4        | 425    | 9 US-10-175-738-570  | Sequence 570, App  |
| 9          | 41    | 55.4        | 425    | 9 US-10-175-752-570  | Sequence 570, App  |
| 10         | 41    | 55.4        | 425    | 9 US-10-176-482-570  | Sequence 570, App  |
| 11         | 41    | 55.4        | 425    | 9 US-10-176-757-570  | Sequence 570, App  |
| 12         | 41    | 55.4        | 425    | 9 US-10-176-913-570  | Sequence 570, App  |
| 13         | 41    | 55.4        | 425    | 9 US-10-180-552-570  | Sequence 570, App  |
| 14         | 41    | 55.4        | 425    | 9 US-10-180-557-570  | Sequence 570, App  |
| 15         | 41    | 55.4        | 425    | 9 US-10-173-700-570  | Sequence 570, App  |
| 16         | 41    | 55.4        | 425    | 9 US-10-174-572-570  | Sequence 570, App  |
| 17         | 41    | 55.4        | 425    | 9 US-10-174-579-570  | Sequence 570, App  |
| 18         | 41    | 55.4        | 425    | 9 US-10-174-582-570  | Sequence 570, App  |
| 19         | 41    | 55.4        | 425    | 9 US-10-174-588-570  | Sequence 570, App  |

|    |    |      |     |                     |                   |
|----|----|------|-----|---------------------|-------------------|
| 20 | 41 | 55.4 | 425 | 9 US-10-175-739-570 | Sequence 570, App |
| 21 | 41 | 55.4 | 425 | 9 US-10-175-740-570 | Sequence 570, App |
| 22 | 41 | 55.4 | 425 | 9 US-10-175-743-570 | Sequence 570, App |
| 23 | 41 | 55.4 | 425 | 9 US-10-176-488-570 | Sequence 570, App |
| 24 | 41 | 55.4 | 425 | 9 US-10-176-492-570 | Sequence 570, App |
| 25 | 41 | 55.4 | 425 | 9 US-10-176-747-570 | Sequence 570, App |
| 26 | 41 | 55.4 | 425 | 9 US-10-176-750-570 | Sequence 570, App |
| 27 | 41 | 55.4 | 425 | 9 US-10-176-985-570 | Sequence 570, App |
| 28 | 41 | 55.4 | 425 | 9 US-10-176-987-570 | Sequence 570, App |
| 29 | 41 | 55.4 | 425 | 9 US-10-176-991-570 | Sequence 570, App |
| 30 | 41 | 55.4 | 425 | 9 US-10-176-992-570 | Sequence 570, App |
| 31 | 41 | 55.4 | 425 | 9 US-10-176-993-570 | Sequence 570, App |
| 32 | 41 | 55.4 | 425 | 9 US-10-184-658-570 | Sequence 570, App |
| 33 | 41 | 55.4 | 425 | 9 US-10-173-695-570 | Sequence 570, App |
| 34 | 41 | 55.4 | 425 | 9 US-10-173-697-570 | Sequence 570, App |
| 35 | 41 | 55.4 | 425 | 9 US-10-173-705-570 | Sequence 570, App |
| 36 | 41 | 55.4 | 425 | 9 US-10-174-585-570 | Sequence 570, App |
| 37 | 41 | 55.4 | 425 | 9 US-10-174-586-570 | Sequence 570, App |
| 38 | 41 | 55.4 | 425 | 9 US-10-174-585-570 | Sequence 570, App |
| 39 | 41 | 55.4 | 425 | 9 US-10-175-747-570 | Sequence 570, App |
| 40 | 41 | 55.4 | 425 | 9 US-10-176-481-570 | Sequence 570, App |
| 41 | 41 | 55.4 | 425 | 9 US-10-176-485-570 | Sequence 570, App |
| 42 | 41 | 55.4 | 425 | 9 US-10-176-487-570 | Sequence 570, App |
| 43 | 41 | 55.4 | 425 | 9 US-10-176-493-570 | Sequence 570, App |
| 44 | 41 | 55.4 | 425 | 9 US-10-176-756-570 | Sequence 570, App |
| 45 | 41 | 55.4 | 425 | 9 US-10-176-911-570 | Sequence 570, App |

## ALIGNMENTS

RESULT 1  
US-09-757-982-5  
Sequence 5, Application US/09757982  
Patent No. US20020094559A1  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: MNI-050  
CURRENT APPLICATION NUMBER: US/09/757,982  
PRIOR FILING DATE: 2001-01-10  
PRIOR APPLICATION NUMBER: 09/163,115  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-757-982-5

Query Match 100.0%; Score 74; DB 10; Length 455;  
Best Local Similarity 100.0%; Pred. No. 0.00097;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KQNSKTSKRGGK 15  
DB 407 KQNSKTSKRGGK 421

RESULT 2  
US-10-081-119-14  
Sequence 14, Application US/10081119  
Publication No. US20030045491A1  
GENERAL INFORMATION:  
APPLICANT: Reinhard, Christoph  
APPLICANT: Jefferson, Anne B.  
APPLICANT: Chan, Vivien W.  
TITLE OF INVENTION: TTK In Diagnosis and as a Therapeutic  
FILE REFERENCE: 16932.002  
CURRENT APPLICATION NUMBER: US/10/081,119  
CURRENT FILING DATE: 2002-02-21



PRIOR APPLICATION NUMBER: 60/289,813  
PRIOR FILING DATE: 2001-02-21  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14  
LENGTH: 841  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-081-119-14

Query Match 60.8%; Score 45; DB 9; Length 841;  
Best Local Similarity 69.2%; Pred. No. 42;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 NSSKTSKRGRGK 15  
Db 829 SSSKTFKKRGRK 841

## RESULT 3

Sequence 260, Application US/09801368  
Patent No. US20020128250A1  
GENERAL INFORMATION:  
APPLICANT: Busby, Robert  
APPLICANT: Cali, Brian  
APPLICANT: Hecht, Peter  
APPLICANT: Holtzman, Doug  
APPLICANT: Madden, Kevin  
APPLICANT: Maxon, Mary  
APPLICANT: Milne, Todd  
APPLICANT: No. US20020128250A1man, Thea  
APPLICANT: Royer, John  
APPLICANT: Salama, Sofie  
APPLICANT: Sherman, Amir  
APPLICANT: Silva, Jeff  
APPLICANT: Summers, Eric  
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
FILE REFERENCE: 109272.147  
CURRENT APPLICATION NUMBER: US/09/801,368  
CURRENT FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 09/487,558  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: US 60/160,587  
PRIOR FILING DATE: 1999-10-20  
NUMBER OF SEQ ID NOS: 440  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 260  
LENGTH: 312  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-801-368-260

Query Match 56.8%; Score 42; DB 10; Length 312;  
Best Local Similarity 64.3%; Pred. No. 45;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 KONSSTTSKRGRK 14  
Db 148 KSNSSPYLNKRGRK 161

## RESULT 4

US-10-174-590-570  
Sequence 570, Application US/10174590  
Publication No. US20030008352A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Matanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C42  
CURRENT APPLICATION NUMBER: US/10/174,590  
CURRENT FILING DATE: 2002-06-18  
Prior application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-174-590-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 KONSSTTSKRGRK 15  
Db 243 KEKASKTTVRRRKR 257

## RESULT 5

US-10-176-758-570  
Sequence 570, Application US/10176758  
Publication No. US20030008353A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Matanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C104  
CURRENT APPLICATION NUMBER: US/10/176,758  
CURRENT FILING DATE: 2002-06-21  
Prior application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-176-758-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 KONSSTTSKRGRK 15  
Db 243 KEKASKTTVRRRKR 257

## RESULT 6

US-10-175-737-570  
Sequence 570, Application US/10175737  
Publication No. US20030013153A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C50  
CURRENT APPLICATION NUMBER: US/10/175,737  
CURRENT FILING DATE: 2002-06-19  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-175-737-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
Db 243 KEKASKTTVRRRKR 257

RESULT 7  
US-10-173-706-570  
Sequence 570, Application US/10173706  
Publication No. US20030022293A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C7  
CURRENT APPLICATION NUMBER: US/10/173,706  
CURRENT FILING DATE: 2002-06-17  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-173-706-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
Db 243 KEKASKTTVRRRKR 257

RESULT 8  
US-10-175-738-570  
Sequence 570, Application US/10175738  
Publication No. US20030022294A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C45  
CURRENT APPLICATION NUMBER: US/10/175,738  
CURRENT FILING DATE: 2002-06-19  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-175-738-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
Db 243 KEKASKTTVRRRKR 257

RESULT 9  
US-10-175-752-570  
Sequence 570, Application US/10175752  
Publication No. US20030022295A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C60  
CURRENT APPLICATION NUMBER: US/10/175,752  
CURRENT FILING DATE: 2002-06-19  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-175-752-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
Db 243 KEKASKTTVRRRKR 257

RESULT 10  
US-10-176-482-570

```
; Sequence 570, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 570
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-176-482-570
```

```
Query Match      55.4%; Score 41; DB 9; Length 425;
Best Local Similarity 53.3%; Pred. No. 86;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
OY      1 KONSSTTSKRRGKK 15
DB      243 KKKSKTTTVRRRRK 257

RESULT 11
US-10-176-757-570
; Sequence 570, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C86
; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 570
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-176-757-570
```

```
Query Match      55.4%; Score 41; DB 9; Length 425;
Best Local Similarity 53.3%; Pred. No. 86;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
RESULT 12
US-10-176-913-570
; Sequence 570, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C66
; CURRENT APPLICATION NUMBER: US/10/176,913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 570
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-176-913-570
```

```
Query Match      55.4%; Score 41; DB 9; Length 425;
Best Local Similarity 53.3%; Pred. No. 86;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
OY      1 KONSSTTSKRRGKK 15
DB      243 KKKSKTTTVRRRRK 257

RESULT 13
US-10-180-552-570
; Sequence 570, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C153
; CURRENT APPLICATION NUMBER: US/10/180,552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 570
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-180-552-570
```

```
Query Match      55.4%; Score 41; DB 9; Length 425;
Best Local Similarity 53.3%; Pred. No. 86;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

Db 243 KEKASKTTTVRRRK 257

RESULT 14

US-10-180-557-570  
; Sequence 570, Application US/10180557  
; Publication No. US20030022301A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C147  
CURRENT APPLICATION NUMBER: US/10/180,557  
CURRENT FILING DATE: 2002-06-25  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-180-557-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 1 KONSSTTSKRGKK 15  
Db 243 KEKASKTTTVRRRK 257

RESULT 15

US-10-173-700-570  
; Sequence 570, Application US/10173700  
; Publication No. US20030027262A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C14  
CURRENT APPLICATION NUMBER: US/10/173,700  
CURRENT FILING DATE: 2002-06-17  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-173-700-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;

Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 1 KONSSTTSKRGKK 15  
Db 243 KEKASKTTTVRRRK 257

Search completed: May 1, 2003, 20:52:15  
Job time: 1.96234 secs





GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:33:12 ; Search time 1.00418 Seconds  
(without alignments)  
439.505 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_407\_421  
Perfect score: 74  
Sequence: 1 KONSSTSKRRGKK 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Archived: 262574 seqs, 29422922 residues  
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfilee1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID                | Description        |
|------------|-------|-------------|--------|----|-------------------|--------------------|
| 1          | 74    | 100.0       | 455    | 3  | US-09-221-235-5   | Sequence 5, Appli  |
| 2          | 74    | 100.0       | 455    | 3  | US-09-221-928-5   | Sequence 5, Appli  |
| 3          | 74    | 100.0       | 455    | 4  | US-09-221-527-5   | Sequence 5, Appli  |
| 4          | 74    | 100.0       | 455    | 4  | US-09-221-236-5   | Sequence 5, Appli  |
| 5          | 74    | 100.0       | 455    | 4  | US-09-221-416-5   | Sequence 5, Appli  |
| 6          | 74    | 100.0       | 455    | 4  | US-09-221-245-5   | Sequence 5, Appli  |
| 7          | 74    | 100.0       | 455    | 4  | US-09-163-115-5   | Sequence 5, Appli  |
| 8          | 74    | 100.0       | 455    | 4  | US-09-221-528-5   | Sequence 5, Appli  |
| 9          | 74    | 100.0       | 455    | 4  | US-09-593-553-5   | Sequence 5, Appli  |
| 10         | 74    | 100.0       | 455    | 4  | US-09-221-237-5   | Sequence 5, Appli  |
| 11         | 39    | 52.7        | 673    | 4  | US-09-078-347A-2  | Sequence 2, Appli  |
| 12         | 38    | 51.4        | 855    | 4  | US-07-956-483-14  | Sequence 14, Appli |
| 13         | 37    | 50.0        | 30     | 2  | US-08-248-839C-80 | Sequence 80, Appli |
| 14         | 37    | 50.0        | 211    | 2  | US-08-708-958-2   | Sequence 2, Appli  |
| 15         | 37    | 50.0        | 357    | 1  | US-08-411-777-9   | Sequence 9, Appli  |
| 16         | 37    | 50.0        | 357    | 3  | US-09-057-088-9   | Sequence 9, Appli  |
| 17         | 37    | 50.0        | 376    | 1  | US-08-090-013-4   | Sequence 4, Appli  |
| 18         | 37    | 50.0        | 376    | 1  | US-08-081-328-4   | Sequence 4, Appli  |
| 19         | 37    | 50.0        | 376    | 1  | US-08-232-249-4   | Sequence 4, Appli  |
| 20         | 37    | 50.0        | 376    | 2  | US-08-833-642A-4  | Sequence 4, Appli  |
| 21         | 37    | 50.0        | 376    | 2  | US-08-389-423-4   | Sequence 4, Appli  |
| 22         | 37    | 50.0        | 376    | 4  | US-09-189-028-4   | Sequence 4, Appli  |
| 23         | 37    | 50.0        | 823    | 4  | US-09-651-656-19  | Sequence 19, Appli |
| 24         | 37    | 50.0        | 823    | 4  | US-09-650-855-19  | Sequence 19, Appli |
| 25         | 36    | 48.6        | 324    | 1  | US-08-323-531-4   | Sequence 4, Appli  |
| 26         | 36    | 48.6        | 324    | 1  | US-08-198-094-4   | Sequence 4, Appli  |
| 27         | 36    | 48.6        | 324    | 4  | US-08-107-794A-4  | Sequence 4, Appli  |

|    |      |      |      |   |                     |                   |
|----|------|------|------|---|---------------------|-------------------|
| 28 | 36   | 48.6 | 324  | 5 | PCT-US93-07424-4    | Sequence 4, Appli |
| 29 | 36   | 48.6 | 324  | 5 | PCT-US95-02087-4    | Sequence 4, Appli |
| 30 | 36   | 48.6 | 401  | 4 | US-09-651-656-103   | Sequence 103, App |
| 31 | 36   | 48.6 | 401  | 4 | US-09-650-855-103   | Sequence 103, App |
| 32 | 36   | 48.6 | 418  | 2 | US-08-978-182-5     | Sequence 5, Appli |
| 33 | 36   | 48.6 | 418  | 2 | US-09-205-681-5     | Sequence 5, Appli |
| 34 | 36   | 48.6 | 602  | 4 | US-09-134-001C-4416 | Sequence 4416, Ap |
| 35 | 36   | 48.6 | 670  | 1 | US-08-363-208-2     | Sequence 2, Appli |
| 36 | 36   | 48.6 | 670  | 4 | US-09-137-478-2     | Sequence 2, Appli |
| 37 | 36   | 48.6 | 671  | 2 | US-08-426-125-8     | Sequence 8, Appli |
| 38 | 36   | 48.6 | 671  | 2 | US-08-455-355-8     | Sequence 8, Appli |
| 39 | 35.5 | 48.0 | 934  | 3 | US-08-929-329-6     | Sequence 6, Appli |
| 40 | 35   | 47.3 | 118  | 3 | US-09-012-084-4     | Sequence 4, Appli |
| 41 | 35   | 47.3 | 304  | 4 | US-09-134-001C-4578 | Sequence 4578, Ap |
| 42 | 35   | 47.3 | 338  | 2 | US-09-047-026A-2    | Sequence 2, Appli |
| 43 | 35   | 47.3 | 534  | 4 | US-09-198-603C-6    | Sequence 6, Appli |
| 44 | 35   | 47.3 | 798  | 4 | US-09-134-001C-4917 | Sequence 4917, Ap |
| 45 | 35   | 47.3 | 1239 | 2 | US-08-937-931-2     | Sequence 2, Appli |

ALIGNMENTS

RESULT 1  
US-09-221-235-5  
; Sequence 5, Application US/09221235  
; Patent No. 6043040  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,235  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-221-235-5

Query Match 100.0%; Score 74; DB 3; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTSKRRGKK 15  
Db 407 KONSSTSKRRGKK 421

RESULT 2  
US-09-221-928-5  
; Sequence 5, Application US/09221928  
; Patent No. 6121030  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,928  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-221-928-5

Query Match 100.0%; Score 74; DB 3; Length 455;

Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
Db 407 KONSSTTSKRRGKK 421

## RESULT 3

US-09-221-527-5  
; Sequence 5, Application US/09221527  
; Patent No. 6146832

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050

; CURRENT APPLICATION NUMBER: US/09/221,527

; CURRENT FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: 09/163,115

; EARLIER FILING DATE:

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 455

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-221-527-5

Query Match 100.0%; Score 74; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
Db 407 KONSSTTSKRRGKK 421

## RESULT 4

US-09-221-236-5  
; Sequence 5, Application US/09221236  
; Patent No. 6146841

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050

; CURRENT APPLICATION NUMBER: US/09/221,236

; CURRENT FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: 09/163,115

; EARLIER FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 455

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-221-236-5

Query Match 100.0%; Score 74; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
Db 407 KONSSTTSKRRGKK 421

## RESULT 5

US-09-221-416-5  
; Sequence 5, Application US/09221416  
; Patent No. 6153417

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,416  
; CURRENT FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: 09/163,115

; EARLIER FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 455

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-221-416-5

Query Match 100.0%; Score 74; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
Db 407 KONSSTTSKRRGKK 421

## RESULT 6

US-09-221-245-5  
; Sequence 5, Application US/09221245  
; Patent No. 6180358

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050

; CURRENT APPLICATION NUMBER: US/09/221,245

; CURRENT FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: US 09/163,115

; EARLIER FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 455

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-221-245-5

Query Match 100.0%; Score 74; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
Db 407 KONSSTTSKRRGKK 421

## RESULT 7

US-09-163-115-5  
; Sequence 5, Application US/09163115A  
; Patent No. 6183962

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050

; CURRENT APPLICATION NUMBER: US/09/163,115A

; CURRENT FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 455

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-163-115-5

Query Match 100.0%; Score 74; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KÖNSSKTTSKRRGKK 15  
Db 407 KÖNSSKTTSKRRGKK 421

## RESULT 8

US-09-221-528-5  
; Sequence 5, Application US/09221528  
; Patent No. 6190874  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,528  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-528-5

Query Match 100.0%; Score 74; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KÖNSSKTTSKRRGKK 15  
Db 407 KÖNSSKTTSKRRGKK 421

## RESULT 9

US-09-593-553-5  
; Sequence 5, Application US/09593553  
; Patent No. 6200770  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/593,553  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: 09/163,115  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-593-553-5

Query Match 100.0%; Score 74; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KÖNSSKTTSKRRGKK 15  
Db 407 KÖNSSKTTSKRRGKK 421

## RESULT 10

US-09-221-237-5  
; Sequence 5, Application US/09221237  
; Patent No. 6214597  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,237  
; CURRENT FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-237-5

Query Match 100.0%; Score 74; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KÖNSSKTTSKRRGKK 15  
Db 407 KÖNSSKTTSKRRGKK 421

## RESULT 11

US-09-078-347A-2  
; Sequence 2, Application US/09078347A  
; Patent No. 6132968  
; GENERAL INFORMATION:  
; APPLICANT: Le, Xiao-Chun  
; APPLICANT: Weinfield, Michael  
; APPLICANT: Xing, James Z.  
; TITLE OF INVENTION: Methods for Quantitating Low Level  
; TITLE OF INVENTION: Modifications of Nucleotide Sequences  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94104

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/078,347A  
; FILING DATE: 13-MAY-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: UALB-03283  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 673 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-09-078-347A-2

Query Match 52.7%; Score 39; DB 4; Length 673;  
Best Local Similarity 61.5%; Pred. No. 56;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ÖNSSKTTSKRRGK 14  
Db 601 ÖNTAKTKAKGRGK 613

## RESULT 12

US-07-956-483-14



```

;
; COMPUTER: IBM Compatible
;
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,839C
; FILING DATE: 25-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valeta A.
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 3614.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
;
US-08-248-839C-80

Query Match 50.0%; Score 37; DB 2; length 30;
Best Local Similarity 72.7%; Pred. No. 4.3;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KNSGKTSKR 11
|||:|||||
Db 10 KNTSKTESLR 20

```

US-08-708-958-2

Query Match  
Best Local Similarity 50.0%; Score 37; DB 2; Length 211;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 SSKTTSKRRGKK 15  
| : : : :  
Db 88 SSSSSSSKRRGKK 99

## RESULT 15

US-08-411-777-9

/ Sequence 9, Application US/08411777  
/ Patent No. 5792641

## / GENERAL INFORMATION:

/ APPLICANT: Schuelein, Martin  
/ APPLICANT: Fredholm, Henrik  
/ APPLICANT: Hjorth, Carsten  
/ APPLICANT: Rasmussen, Grethe  
/ APPLICANT: Nielsen, Egon  
/ APPLICANT: Rosholm, Peter  
/ TITLE OF INVENTION: Cellulase Variants  
/ NUMBER OF SEQUENCES: 10  
/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: No. 57926410 No. 5792641disk of No. 5792641th America  
/ STREET: 405 Lexington Avenue  
/ CITY: New York  
/ STATE: NY

/ COUNTRY: USA  
/ ZIP: 10174-6401

/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Diskette

/ COMPUTER: IBM Compatible  
/ OPERATING SYSTEM: DOS

/ SOFTWARE: FastSeq for Windows Version 2.0  
/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/411,777  
/ FILING DATE: 05-MAY-1995

/ CLASSIFICATION: 435

/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER:

/ FILING DATE:

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Gregg, Valeta A

/ REGISTRATION NUMBER: 35,127

/ REFERENCE/DOCKET NUMBER: 3913.504-US

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: 212-867-0123

/ TELEFAX: 212-868-9655

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: No. 5792641e

/ US-08-411-777-9

## Query Match

Best Local Similarity 50.0%; Score 37; DB 1; Length 357;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 NSSKTSKRRGKK 15  
| : : : :  
Db 289 NKPKTTQKVRGTK 301

Search completed: May 1, 2003, 20:39:43'  
Job time : 3.00418 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:51:28 ; Search time 47 Seconds  
(without alignments)  
930.663 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455

Sequence: 1 MSSUGASFVQIKFDDLPFE.....GDDDDDDGEEEDNDMDNSE 455

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database:

PIR\_73: \*  
1: pirl: \*  
2: pirl2: \*  
3: pirl3: \*  
4: pirl4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 11    | 2.4         | 848    | 2 B87950 | protein F33E2.2 [i |
| 2          | 11    | 2.4         | 855    | 2 T20082 | hypothetical prote |
| 3          | 10    | 2.2         | 729    | 2 A29651 | KEK1 protein precu |
| 4          | 10    | 2.2         | 1089   | 2 S48244 | NMD2 protein - yea |
| 5          | 9     | 2.0         | 297    | 2 T30613 | hypothetical prote |
| 6          | 9     | 2.0         | 375    | 2 A39622 | protein kinase MCK |
| 7          | 9     | 2.0         | 408    | 2 T08069 | protein kinase, 48 |
| 8          | 9     | 2.0         | 502    | 2 JC4337 | activin receptor 1 |
| 9          | 9     | 2.0         | 502    | 2 I48241 | ALK-1 - mouse      |
| 10         | 9     | 2.0         | 764    | 2 S64951 | hypothetical prote |
| 11         | 9     | 2.0         | 856    | 2 T52415 | polycarb protein E |
| 12         | 9     | 2.0         | 898    | 2 T01503 | hypothetical prote |
| 13         | 9     | 2.0         | 1363   | 2 T43220 | insulin-like growt |
| 14         | 9     | 2.0         | 1390   | 2 T30346 | insulin receptor - |
| 15         | 9     | 2.0         | 2391   | 2 T18410 | cardamoyl-phosphat |
| 16         | 8     | 1.8         | 101    | 2 I38220 | protein-serine/thr |
| 17         | 8     | 1.8         | 108    | 2 I38212 | hypothetical prote |
| 18         | 8     | 1.8         | 157    | 2 S09814 | 2S albumin precurs |
| 19         | 8     | 1.8         | 158    | 2 B83130 | hypothetical prote |
| 20         | 8     | 1.8         | 162    | 2 A96520 | hypothetical prote |
| 21         | 8     | 1.8         | 198    | 2 D69413 | conserved hypothet |
| 22         | 8     | 1.8         | 211    | 2 PQBP82 | antiterminator Q - |
| 23         | 8     | 1.8         | 229    | 1 D90831 | antitermination pr |
| 24         | 8     | 1.8         | 229    | 2 D90831 | probable antitermi |
| 25         | 8     | 1.8         | 229    | 2 G85688 | hypothetical prote |
| 26         | 8     | 1.8         | 237    | 2 T47271 | hypothetical prote |
| 27         | 8     | 1.8         | 250    | 2 T47611 | hypothetical prote |
| 28         | 8     | 1.8         | 260    | 2 G84645 | hypothetical prote |
| 29         | 8     | 1.8         | 286    | 2 A34599 | DNA-binding protei |

|     |   |     |      |          |                    |
|-----|---|-----|------|----------|--------------------|
| 30  | 8 | 1.8 | 294  | 1 A37818 | osteopontin precur |
| 31  | 8 | 1.8 | 298  | 2 T52117 | zinc finger protei |
| 32  | 8 | 1.8 | 301  | 2 JC5811 | osteopontin - rat  |
| 33  | 8 | 1.8 | 317  | 1 A25917 | conserved hypothet |
| 34  | 8 | 1.8 | 318  | 2 E69026 | hypothetical prote |
| 35  | 8 | 1.8 | 328  | 2 T16747 | ribosomal protein  |
| 36  | 8 | 1.8 | 352  | 1 R5HSL0 | 50S ribosomal prot |
| 37  | 8 | 1.8 | 352  | 2 G84266 | protein FIN19.7 [i |
| 38  | 8 | 1.8 | 368  | 2 G96668 | hypothetical prote |
| 39  | 8 | 1.8 | 374  | 2 T33173 | hypothetical prote |
| 40  | 8 | 1.8 | 379  | 2 T04645 | hypothetical prote |
| 41  | 8 | 1.8 | 390  | 2 T01451 | protein kinase hom |
| 42  | 8 | 1.8 | 391  | 2 D64366 | hypothetical prote |
| 43  | 8 | 1.8 | 422  | 2 S52578 | serine/threonine-s |
| 44  | 8 | 1.8 | 426  | 2 T05676 | hypothetical prote |
| 45  | 8 | 1.8 | 434  | 2 B54843 | nemo, form II - fr |
| 46  | 8 | 1.8 | 436  | 2 T51402 | serine/threonine-s |
| 47  | 8 | 1.8 | 439  | 2 T31734 | hypothetical prote |
| 48  | 8 | 1.8 | 471  | 2 T39232 | probable serine th |
| 49  | 8 | 1.8 | 477  | 2 A54843 | nemo, form I - fru |
| 50  | 8 | 1.8 | 482  | 2 S37845 | transcription init |
| 51  | 8 | 1.8 | 485  | 2 T49237 | hypothetical prote |
| 52  | 8 | 1.8 | 494  | 2 D84860 | probable protein k |
| 53  | 8 | 1.8 | 504  | 2 T10558 | hypothetical prote |
| 54  | 8 | 1.8 | 513  | 2 B96524 | hypothetical prote |
| 55  | 8 | 1.8 | 526  | 2 T47786 | hypothetical prote |
| 56  | 8 | 1.8 | 538  | 2 C83284 | probable biotin-de |
| 57  | 8 | 1.8 | 541  | 2 A12238 | glutamine-fructose |
| 58  | 8 | 1.8 | 560  | 2 T49839 | hypothetical prote |
| 59  | 8 | 1.8 | 567  | 2 JC5957 | transforming growt |
| 60  | 8 | 1.8 | 579  | 2 JC5955 | transforming growt |
| 61  | 8 | 1.8 | 584  | 2 T49206 | phosphoinositide-s |
| 62  | 8 | 1.8 | 590  | 2 S57594 | hypothetical prote |
| 63  | 8 | 1.8 | 606  | 2 JC5956 | transforming growt |
| 64  | 8 | 1.8 | 684  | 2 C96596 | hypothetical prote |
| 65  | 8 | 1.8 | 690  | 2 C96572 | protein F12M16.4 [ |
| 66  | 8 | 1.8 | 746  | 2 G02838 | enhancer-of-zeste  |
| 67  | 8 | 1.8 | 800  | 2 T02852 | probable membrane  |
| 68  | 8 | 1.8 | 836  | 2 B96716 | probable serine/th |
| 69  | 8 | 1.8 | 840  | 2 S48975 | hypothetical prote |
| 70  | 8 | 1.8 | 847  | 1 A53800 | mixed-lineage prot |
| 71  | 8 | 1.8 | 872  | 2 T18861 | probable peptide-a |
| 72  | 8 | 1.8 | 938  | 2 A56731 | chromatin assembly |
| 73  | 8 | 1.8 | 943  | 2 B45082 | neurotrophic recep |
| 74  | 8 | 1.8 | 992  | 2 T05335 | hypothetical prote |
| 75  | 8 | 1.8 | 992  | 2 S49835 | hypothetical prote |
| 76  | 8 | 1.8 | 1019 | 2 T13039 | tyrosine kinase re |
| 77  | 8 | 1.8 | 1038 | 2 T02634 | rep protein homolo |
| 78  | 8 | 1.8 | 1085 | 2 S55352 | IFH1 protein - yea |
| 79  | 8 | 1.8 | 1128 | 2 G86266 | hypothetical prote |
| 80  | 8 | 1.8 | 1131 | 2 S22266 | FUN30 protein - ye |
| 81  | 8 | 1.8 | 1165 | 2 S62982 | vacuolar protein V |
| 82  | 8 | 1.8 | 1234 | 2 T00363 | hypothetical prote |
| 83  | 8 | 1.8 | 1341 | 2 S66835 | probable membrane  |
| 84  | 8 | 1.8 | 2251 | 2 T24490 | hypothetical prote |
| 85  | 8 | 1.8 | 3394 | 2 T18501 | hypothetical prote |
| 86  | 7 | 1.5 | 51   | 2 PN0479 | cell division cycl |
| 87  | 7 | 1.5 | 54   | 2 S56724 | abscisic acid-indu |
| 88  | 7 | 1.5 | 54   | 2 S56725 | abscisic acid-indu |
| 89  | 7 | 1.5 | 71   | 2 A69154 | conserved hypothet |
| 90  | 7 | 1.5 | 85   | 2 F97131 | hypothetical prote |
| 91  | 7 | 1.5 | 95   | 2 T46086 | hypothetical prote |
| 92  | 7 | 1.5 | 101  | 1 RGECKK | transcription repr |
| 93  | 7 | 1.5 | 107  | 2 T35523 | probable small sec |
| 94  | 7 | 1.5 | 108  | 2 T16893 | hypothetical prote |
| 95  | 7 | 1.5 | 111  | 2 T38218 | protein-serine/thr |
| 96  | 7 | 1.5 | 114  | 2 D46516 | Ig lambda chain V  |
| 97  | 7 | 1.5 | 116  | 2 AD2693 | conserved hypothet |
| 98  | 7 | 1.5 | 123  | 2 T49794 | hypothetical prote |
| 99  | 7 | 1.5 | 133  | 2 T48978 | hypothetical prote |
| 100 | 7 | 1.5 | 133  | 2 AE1435 | PTS system, fructo |



## ALIGNMENTS

## RESULT 1

B87950  
Protein F33E2.2 [imported] - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C;Accession: B87950  
R;anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A;Reference number: A75000; MUID:99069613; PMID:9851916  
A;Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A;Accession: B87950  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-848 <STO>  
A;Cross-references: GB:chr\_I; PIDN:CAB06544.1; PID:g3876653; GSPDB:GN00019; CESP:F33E2.2  
C;Genetics:  
Gene: F33E2.2  
Map position: 1

Query Match 2.4%; Score 11; DB 2; Length 848;  
Best Local Similarity 100.0%; Pred. No. 0.044;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 193 GVLWEMLTRE 203  
|||  
Db 235 GVLWEMLTRE 245

## RESULT 2

T20082  
hypothetical protein F33E2.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C;Accession: T20082; T21703  
R;Lennard, N.  
submitted to the EMBL Data Library, April 1998  
A;Reference number: Z19221  
A;Accession: T20082  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-855 <WIL>  
A;Cross-references: EMBL:AL022593; PIDN:CAA18635.2; GSPDB:GN00019; CESP:F33E2.2  
A;Experimental source: clone C49G9  
Lennard, N.  
mitted to the EMBL Data Library, January 1997  
A;Reference number: Z19461  
A;Accession: T21703  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-855 <WIL>  
A;Cross-references: EMBL:Z84574; PIDN:CAB06544.2; GSPDB:GN00019; CESP:F33E2.2  
A;Experimental source: clone F33E2  
C;Genetics:  
A;Gene: CESP:F33E2.2  
A;Map position: 1  
A;Introns: 47/2; 213/2; 263/2; 366/3; 494/1; 547/3; 608/1; 826/3

Query Match 2.4%; Score 11; DB 2; Length 855;  
Best Local Similarity 100.0%; Pred. No. 0.044;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 193 GVLWEMLTRE 203  
|||  
Db 235 GVLWEMLTRE 245

RESULT 3  
A29651

KEK1 protein precursor - yeast (Saccharomyces cerevisiae)  
N;Alternate names: protein G1237; protein YGL203C  
C;Species: Saccharomyces cerevisiae  
C;Date: 08-Mar-1989 #sequence\_revision 08-Mar-1989 #text\_change 21-Jul-2000  
C;Accession: A29651; S64221  
R;Dmochowska, A.; Dignard, D.; Henning, D.; Thomas, D.Y.; Bussey, H.  
Cell 50, 573-584, 1987  
A;Title: Yeast KEK1 gene encodes a putative protease with a carboxypeptidase B-like fun  
A;Reference number: A29651; MUID:87273520; PMID:3301004  
A;Accession: A29651  
A;Molecule type: DNA  
A;Residues: 1-729 <DMO>  
A;Cross-references: GB:M17231; NID:g171778; PIDN:AAA34717.1; PID:g171779  
R;Bourson, A.J.; McReynolds, A.D.K.; Wright, L.F.  
submitted to the Protein Sequence Database, May 1996  
A;Reference number: S64218  
A;Accession: S64221  
A;Molecule type: DNA  
A;Residues: 1-729 <BJO>  
A;Cross-references: EMBL:Z72725; NID:g1322835; PID:e243802; PID:g1322836; MIPS:YGL203C  
A;Experimental source: strain S288C  
C;Genetics:  
A;Gene: SGD:KEK1  
A;Cross-references: SGD:S0003171; MIPS:YGL203C  
A;Map position: 7L  
C;Keywords: Golgi apparatus; transmembrane protein  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-729/Product: KEK1 protein #status predicted <MAT>  
F;619-635/Domain: transmembrane #status predicted <TMM>

Query Match 2.2%; Score 10; DB 2; Length 729;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 434 SEGDDDDDD 443  
|||  
Db 564 SEGDDDDDD 573

## RESULT 4

S48244  
NMD2 protein - yeast (Saccharomyces cerevisiae)  
N;Alternate names: IFS1 protein; protein YHR077C  
C;Species: Saccharomyces cerevisiae  
C;Date: 28-Oct-1994 #sequence\_revision 10-Feb-1995 #text\_change 23-Mar-2001  
C;Accession: S48244; S46815; S64648; S64738  
R;He, F.; Jacobson, A.  
submitted to the EMBL Data Library, September 1994  
A;Description: Identification of a novel component of the nonsense-mediated mRNA decay  
A;Reference number: S48244  
A;Accession: S48244  
A;Molecule type: DNA  
A;Residues: 1-1089 <HEF>  
A;Cross-references: EMBL:U14974; NID:g555938; PID:g555939  
R;Favell, T.  
submitted to the EMBL Data Library, June 1994  
A;Description: The sequence of S. cerevisiae cosmid 9205.  
A;Reference number: S46795  
A;Accession: S46815  
A;Molecule type: DNA  
A;Residues: 'MYOO', 3-1089 <FAV>  
A;Cross-references: EMBL:U10556; NID:g500825; PID:g500836; MIPS:YHR077C  
R;Lee, S.I.; Umen, J.G.; Varmus, H.E.  
Proc. Natl. Acad. Sci. U.S.A. 92, 6587-6591, 1995  
A;Title: A genetic screen identifies cellular factors involved in retroviral -1 framesh  
A;Reference number: S64648; MUID:95327692; PMID:7604038  
A;Accession: S64648  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-1089 <LEE>  
A;Cross-references: EMBL:U28158  
R;Varmus, H.E.  
submitted to the EMBL Data Library, May 1995

A/Reference number: S64738

A/Accession: S64738

A/Molecule type: DNA

A/Residues: 'MYQQ',3-1089 <VAR>

A/Cross-references: EMBL:U28158; NID:g967212; PID:g967213

C/Genetics:

A/Gene: SGD:NMD2; IFS1

A/Cross-references: SGD:S0001119; MIPS:YHR077C

A/Map position: 8R

A/Introns: 2/3

#### Query Match

Best Local Similarity 2.0%; Score 10; DB 2; Length 1089;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446

Db 902 DDDDDDDGEE 911

#### RESULT 5

T30613

hypothetical protein 11L - Molluscum contagiosum virus 1

N/Alternate names: MC011L

C/Species: Molluscum contagiosum virus 1

C/Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 11-May-2000

C/Accession: T30613

R/Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.

Science 273, 813-816, 1996

A/Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re

A/Reference number: 220876; MUID:96325459; PMID:8670425

A/Accession: T30613

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-297 <SEN>

A/Cross-references: EMBL:U60315; PIDN:AAC55139.1

C/Genetics:

A/Note: MC011L

#### Query Match

Best Local Similarity 2.0%; Score 9; DB 2; Length 297;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDDG 444

Db 206 GDDDDDDDG 214

#### RESULT 6

A39622

protein kinase MCK1 (EC 2.7.1.-) - Yeast (Saccharomyces cerevisiae)

N/Alternate names: protein kinase YPK1; protein N0392; protein YNL307C

C/Species: Saccharomyces cerevisiae

C/Date: 24-Jan-1992 #sequence\_revision 24-Jan-1992 #text\_change 23-Mar-2001

C/Accession: A39622; A39623; A36354; B36354; S51304; S59573; S63288; S63283

R/Neugeborn, L.; Mitchell, A.P.

Genes Dev. 5, 533-548, 1991

A/Title: The yeast MCK1 gene encodes a protein kinase homolog that activates early meiot

A/Reference number: A39622; MUID:91184610; PMID:2010083

A/Accession: A39622

A/Molecule type: DNA

A/Residues: 1-375 <NEI>

A/Cross-references: GB:X55054; NID:g3907; PIDN:CAA38895.1; PID:g3908

R/Shero, J.H.; Hietter, P.

Genes Dev. 5, 549-560, 1991

A/Title: A suppressor of a centromere DNA mutation encodes a putative protein kinase (MC

A/Reference number: A39623; MUID:91184611; PMID:2010084

A/Accession: A39623

A/Molecule type: DNA

A/Residues: 1-375 <SHE>

A/Cross-references: GB:M55984; NID:g171909; PIDN:AAA34764.1; PID:g171910

R/Dailey, D.; Schieven, G.L.; Lim, M.Y.; Marguardt, H.; Gilmore, T.; Thorner, J.; Martin

Mol. Cell. Biol. 10, 6244-6256, 1990

A/Title: Novel yeast protein kinase (YPK1 gene product) is a 40-kilodalton phosphotyros

A/Reference number: A36354; MUID:91061730; PMID:1701015

A/Accession: A36354

A/Molecule type: DNA

A/Residues: 1-375 <DAI>

A/Cross-references: GB:M55984; NID:g171909; PIDN:AAA34764.1; PID:g171910

A/Accession: B36354

A/Molecule type: protein

A/Residues: 41-58;70-94;96-108;134-139;182-188;238-254;262-306 <DA2>

R/Nicaud, J.J.

submitted to the EMBL Data Library, January 1995

A/Description: Sequence analysis of a 13.9 kb fragment of yeast chromosome XIV identifi

A/Reference number: S51285

A/Accession: S51304

A/Molecule type: DNA

A/Residues: 1-375 <NIC>

A/Cross-references: EMBL:Z46259; NID:g633655; PIDN:CAA86388.1; PID:g633675

R/Maftah, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.

Yeast 11, 1077-1085, 1995

A/Title: Sequencing analysis of a 24.7 kb fragment of yeast chromosome XIV identifies s

A/Reference number: S59562; MUID:96076632; PMID:7502583

A/Accession: S59573

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-375 <MAF>

A/Cross-references: EMBL:Z46259; NID:g633655; PIDN:CAA86388.1; PID:g633675

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995

R/Maftah, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.

submitted to the Protein Sequence Database, April 1996

A/Reference number: S63287

A/Accession: S63288

A/Molecule type: DNA

A/Residues: 1-375 <MAW>

A/Cross-references: EMBL:Z71583; NID:g1302406; PIDN:CAA96236.1; PID:g1302407; MIPS:YNL3

A/Experimental source: strain S288C

R/Maurex, C.T.C.; Urbanus, J.H.M.; Planta, R.J.

submitted to the Protein Sequence Database, April 1996

A/Reference number: S63266

A/Accession: S63283

A/Molecule type: DNA

A/Residues: 1-17 <MAU>

A/Cross-references: EMBL:Z71583; MIPS:YNL307C

A/Experimental source: strain S288C

C/Genetics:

A/Gene: SGD:MCK1; YPK1

A/Cross-references: SGD:S0005251; MIPS:YNL307C

A/Map position: 14L

C/Function:

A/Description: phosphotransferase

C/Superfamily: kinase-related transforming protein; protein kinase homology

C/Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein

F/33-302/Domain: protein kinase homology <KIN>

F/41-49/Region: protein kinase ATP-binding motif

#### Query Match

Best Local Similarity 2.0%; Score 9; DB 2; Length 375;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 GVLKICDFG 153

Db 177 GVLKICDFG 185

#### RESULT 7

T08069

protein kinase, 48K - Chlamydomonas reinhardtii

C/Species: Chlamydomonas reinhardtii

C/Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 08-Oct-1999

C/Accession: T08069

R/Kutvari, V.; Zhang, Y.; Luo, Y.; Snell, W.J.

Proc. Natl. Acad. Sci. U.S.A. 93, 39-43, 1996

A/Title: Molecular cloning of a protein kinase whose phosphorylation is regulated by ger

A/Reference number: Z16331; MUID:96133873; PMID:8552645

A/Accession: T08069  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-408 <KUR>  
A/Cross-references: EMBL:U36196; NID:g1019886; PIDN:AAA6956.1; PID:g1019887  
A/Experimental source: strain 21gr(+)  
A/Function:  
A/Description: may be involved in a signalling pathway in fertilization  
C/Keywords: autophosphorylation

Query Match 2.0%; Score 9; DB 2; Length 408;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDDEDDDD 443  
|||||  
Db 393 EGDDEDDDD 401

## RESULT 8

CA4337

Activin receptor like kinase-1 homolog - mouse  
N/Alternate names: activin type IB receptor

C/Species: Mus musculus (house mouse)

C/Date: 06-Dec-1995 #sequence\_revision 08-Feb-1996 #text\_change 21-Jul-2000

C/Accession: J04337; I49478

R/Wu, X.; Robinson, C.E.; Fong, H.W.; Crabtree, J.S.; Rodriguez, B.R.; Roe, B.A.; Gimble  
Biochem. Biophys. Res. Commun. 216, 78-83, 1995

A/Title: Cloning and characterization of the murine activin receptor like kinase-1 (ALK-  
A/Reference number: J04337; MUID:96067535; PMID:7488127

A/Accession: J04337

A/Molecule type: mRNA

A/Residues: 1-502 <WUX>

A/Cross-references: GB:I48015; NID:g1408066; PIDN:AAB03642.1; PID:g1020393

A/Experimental source: lung

R/Dewulf, N.; Verschuere, K.; Lonnoy, O.; Moren, A.; Grimsby, S.; Vande Spiegle, K.; Mi  
Endocrinology 136, 2652-2663, 1995

A/Title: Distinct spatial and temporal expression patterns of two type I receptors for b  
A/Reference number: I48241; MUID:95269711; PMID:7750489

A/Accession: I49478

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-502 <RES>

A/Cross-references: GB:I48015; NID:g1408066; PIDN:AAB03642.1; PID:g1020393

C/Comment: This protein belongs to transforming growth factor beta family, and it is a c  
C/Genetics:

A/Gene: ALK-1

C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C/Keywords: ATP; glycoprotein; receptor; transmembrane protein

F:118-139/Domain: transmembrane #status predicted <TMM>

F:199-495/Domain: protein kinase homology <KIN>

F:207-215/Region: protein kinase ATP-binding motif

F:97/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 2.0%; Score 9; DB 2; Length 502;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
|||||  
Db 327 HRDLKSRNV 335

## RESULT 9

I48241

ALK-1 - mouse

C/Species: Mus musculus (house mouse)

C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 24-Sep-1999

C/Accession: I48241

R/Dewulf, N.; Verschuere, K.; Lonnoy, O.; Moren, A.; Grimsby, S.; Vande Spiegle, K.; Mi  
Endocrinology 136, 2652-2663, 1995

A/Title: Distinct spatial and temporal expression patterns of two type I receptors for b  
A/Reference number: I48241; MUID:95269711; PMID:7750489

A/Accession: I48241  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-502 <RES>  
A/Cross-references: EMBL:Z31664; NID:g840814; PIDN:CAA83484.1; PID:g840815  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C/Keywords: ATP  
F:199-495/Domain: protein kinase homology <KIN>  
F:207-215/Region: protein kinase ATP-binding motif

Query Match 2.0%; Score 9; DB 2; Length 502;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
|||||  
Db 327 HRDLKSRNV 335

## RESULT 10

S64951

hypothetical protein YLR114C - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein L2941

C/Species: Saccharomyces cerevisiae

C/Date: 01-Aug-1995 #sequence\_revision 24-May-1996 #text\_change 19-Apr-2002

C/Accession: S64951; S69401

R/Verhaesselt, P.; Voet, M.; Volckaert, G.  
submitted to the protein Sequence Database, May 1996

A/Reference number: S64943

A/Accession: S64951

A/Molecule type: DNA

A/Residues: 1-764 <VER>

A/Cross-references: EMBL:Z73286; NID:g1360509; PIDN:CAA97681.1; PID:e245807; PID:g13605

A/Experimental source: strain S288C

R/Verhaesselt, P.; Volckaert, G.  
submitted to the EMBL Data Library, September 1995

A/Reference number: S69393

A/Accession: S69401

A/Molecule type: DNA

A/Residues: 1-764 <VER>

A/Cross-references: EMBL:X89514; NID:g1297019; PIDN:CAA61692.1; PID:e198747; PID:g12970

C/Genetics:

A/Gene: SGD:EFR4

A/Cross-references: SGD:S0004104

A/Map position: 12R

Query Match 2.0%; Score 9; DB 2; Length 764;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGE 445  
|||||  
Db 628 DDDDDDDGE 636

RESULT 11  
T52415  
polycomb protein E2A1 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 24-Oct-2000

C/Accession: T52415

R/Bilodeau, P.; Luo, M.; Dennis, E.S.; Peacock, W.J.; Chaudhury, A.M.  
submitted to the EMBL Data Library, October 1998

A/Reference number: Z26069

A/Accession: T52415

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-856 <BIL>

A/Cross-references: EMBL:AF100163; PIDN:AAD09108.1

Query Match 2.0%; Score 9; DB 2; Length 856;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 DDDDDGEE 447  
Db 17 DDDDDGEE 25

## RESULT 12

T01503  
hypothetical protein T10M13.3 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 24-Mar-1999  
C/Accession: T01503  
R/Johnson, A.F.; de la Bastide, M.; Lodhi, M.; Hoffman, J.; Hasegawa, A.; Gnoj, L.; Gott  
Martensen, R.; McCombie, W.  
Submitted to the EMBL Data Library, May 1997  
A/Description: The sequence of the Arabidopsis thaliana T10M13 BAC.  
A/Reference number: Z14346  
A/Accession: T01503  
Status: translated from GB/EMBL/DBJ  
Molecule type: DNA  
Residues: 1-898 <JOH>

A/Cross-references: EMBL:AF001308; NID:G2104523; PID:G3912918  
A/Experimental source: cultivar Columbia  
C/Genetics:  
A/Map position: 4S  
A/Introns: 48/2; 83/1; 118/3; 207/2; 260/2; 288/3; 336/3; 352/3; 554/2; 603/3; 677/2; 72  
A/Note: T10M13.3

Query Match  
Best Local Similarity 100.0%; Score 9; DB 2; Length 898;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 DDDDDGEE 447  
Db 83 DDDDDGEE 91

## RESULT 13

T43220  
Insulin-like growth factor-I receptor - common lancelet  
N/Alternate names: Insulin-like peptide receptor  
C/Species: Branchiostoma lanceolatum (common lancelet)  
C/Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 11-May-2000  
C/Accession: T43220  
R/Pashmforoush, M.; Chan, S.J.; Steiner, D.F.  
Mol. Endocrinol. 10, 857-866, 1996  
A/Title: Structure and expression of the insulin-like peptide receptor from amphioxus.  
A/Reference number: Z22346; MUID:96408719; PMID:8813726  
A/Accession: T43220  
Status: preliminary; translated from GB/EMBL/DBJ  
Molecule type: mRNA  
A/Residues: 1-1363 <PAS>

A/Cross-references: EMBL:S83394; NID:G1911771; PID:G1911772; PIDN:AAB50848.1  
C/Superfamily: Insulin receptor; protein kinase homology  
C/Keywords: hormone receptor

Query Match  
Best Local Similarity 100.0%; Score 9; DB 2; Length 1363;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVLWEM 199  
Db 1210 SYGVVLWEM 1218

## RESULT 14

T30346  
Insulin receptor - yellow fever mosquito  
C/Species: Aedes aegypti (yellow fever mosquito)  
C/Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 31-Jan-2000  
C/Accession: T30346  
R/Graf, R.; Neuenschwander, S.; Brown, M.R.; Ackermann, U.  
Insect Mol. Biol. 6, 151-163, 1996

A/Title: Insulin mediated secretion of ecdysteroids from mosquito ovaries and molecular  
A/Reference number: Z20834  
A/Accession: T30346  
Status: preliminary; translated from GB/EMBL/DBJ  
Molecule type: mRNA  
A/Residues: 1-1390 <GRA>  
A/Cross-references: EMBL:U72939; NID:G1620749; PID:G1620750; PIDN:AAB17094.1  
C/Superfamily: Insulin receptor; protein kinase homology

Query Match  
Best Local Similarity 100.0%; Score 9; DB 2; Length 1390;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVLWEM 199  
Db 1224 SYGVVLWEM 1232

## RESULT 15

T18410  
carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5) II - malaria parasite  
C/Species: Plasmodium falciparum  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000  
C/Accession: T18410  
R/Flores, M.V.C.; O'Sullivan, W.J.; Stewart, T.S.  
Submitted to the EMBL Data Library, March 1997  
A/Description: Characterisation of the carbamoyl phosphate synthetase gene from Plasmod  
A/Reference number: Z18931  
A/Accession: T18410  
Status: preliminary; translated from GB/EMBL/DBJ  
Molecule type: DNA  
A/Residues: 1-2391 <FLO>

A/Cross-references: EMBL:L32150; NID:G476023; PID:G476024; PIDN:AA29522.1  
C/Superfamily: rudimentary enzyme; aspartate/ornithine carbamoyltransferase homology; B  
arbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homology; carbamoyl-pho  
C/Keywords: ligase  
F/96-2367/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>

Query Match  
Best Local Similarity 100.0%; Score 9; DB 2; Length 2391;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDGEE 445  
Db 1786 DDDDDGEE 1794

## RESULT 16

I38220  
protein-serine/threonine kinase - human (fragment)  
C/Species: Homo sapiens (man)  
C/Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 24-Sep-1999  
C/Accession: I38220; S37422  
R/Schultz, S.J.; Nigg, E.A.  
Cell Growth Differ. 4, 821-830, 1993  
A/Title: Identification of 21 novel human protein kinases, including 3 members of a fam  
A/Reference number: I38211; MUID:94100173; PMID:8274451  
A/Accession: I38220  
Status: preliminary; translated from GB/EMBL/DBJ  
Molecule type: mRNA  
A/Residues: 1-101 <RES>  
A/Cross-references: EMBL:Z25430; NID:G405738; PIDN:CAA80917.1; PID:G405739  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol  
F/1-101/Domain: protein kinase homology (fragment) <KIN>

Query Match  
Best Local Similarity 100.0%; Score 8; DB 2; Length 101;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
Db 92 WMAPEVIQ 99



RESULT 17  
I38212  
protein-serine/threonine kinase - human (fragment)  
C/Species: Homo sapiens (man)  
C/Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 24-Sep-1999  
C/Accession: I38212; S37414  
R/Schultz, S.J.; Ni99, E.A.  
Cell Growth Differ. 4, 821-830, 1993  
A/Title: Identification of 21 novel human protein kinases, including 3 members of a family  
A/Reference number: I38211; MUID:94100173; PMID:8274451  
A/Accession: I38212  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-108 <RES>  
A/Cross-references: EMBL:Z25422; NID:9405722; PIDN:CAA80909.1; PID:9405723  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
F:1-108/Domain: protein kinase homology (fragment) <KIN>

Query Match 1.8%; Score 8; DB 2; Length 108;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 WMAPEVIO 179  
DB 92 WMAPEVIO 99

RESULT 18  
S09814  
hypothetical protein UL51 - human cytomegalovirus (strain AD169)  
N/Alternate names: hypothetical protein HPLF3  
C/Species: human cytomegalovirus, human herpesvirus 5  
A/Note: host Homo sapiens (man)  
C/Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 08-Oct-1999  
C/Accession: S09814  
R/Chee, M.S.; Bankier, A.T.; Beck, S.; Bohm, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;  
M.; Barrell, B.G.  
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990  
A/Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus  
A/Reference number: S09749; MUID:90269039; PMID:2161319  
A/Accession: S09814  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-157 <CHB>  
A/Cross-references: EMBL:X17403; NID:959591; PIDN:CAA35410.1; PID:91780829  
A/Note: this sequence was submitted to the EMBL Data Library, December 1989  
C/Superfamily: varicella-zoster virus gene 25 protein

Query Match 1.8%; Score 8; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 8.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 440 DDDGEEE 447  
DB 12 DDDGEEE 19

RESULT 19  
T05710  
2S albumin precursor - soybean  
N/Alternate names: aspartic acid-rich peptide  
C/Species: Glycine max (soybean)  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C/Accession: T05710; A28485  
R/Wang, J.; Pichersky, E.  
Plant Physiol. 114, 1567, 1997  
A/Title: Nucleotide sequence of S-Adenosyl-L-Methionine:Caffeic Acid 3-O-Methyltransferase  
A/Reference number: Z15424  
A/Accession: T05710  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-158 <WAN>

A/Cross-references: EMBL:AF005030; NID:92305019; PID:92305020  
A/Experimental source: cultivar Hodgson 78; cotyledon  
R/Odani, S.; Koide, T.; Ono, T.  
J. Biol. Chem. 262, 10502-10505, 1987  
A/Title: Amino acid sequence of a soybean (Glycine max) seed polypeptide having a poly(I)  
A/Reference number: A28485; MUID:87280104; PMID:3611081  
A/Accession: A28485  
A/Molecule type: protein  
A/Residues: 22-64 <ODA>  
C/Superfamily: soybean 2S albumin  
C/Keywords: glycoprotein  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-64/Product: 2S albumin small chain #status predicted <MAT>  
F:54-56/Region: cell attachment (R-G-D) motif  
F:39/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.8%; Score 8; DB 1; Length 158;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443  
DB 55 GDDDDDD 62

RESULT 20  
B83130  
hypothetical protein PA4129 [imported] - Pseudomonas aeruginosa (strain PA01)  
C/Species: Pseudomonas aeruginosa  
C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C/Accession: B83130  
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
A/Reference number: A82950; MUID:20437337; PMID:10984043  
A/Accession: B83130  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-162 <STO>  
A/Cross-references: GB:AE004829; GB:AE004091; NID:99950327; PIDN:AA607516.1; GSPDB:GN00  
C/Genetics:  
A/Gene: PA4129

Query Match 1.8%; Score 8; DB 2; Length 162;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 329 PLLPLPLA 336  
DB 34 PLLPLPLA 41

RESULT 21  
A96520  
hypothetical protein T2J15.12 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: A96520  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: A96520

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-198 <STO>  
A/Cross-references: GB:AE005173; NID:g10645475; PIDN:AAG21589.1; GSPDB:GN00141  
C/Genetics:  
A/Gene: T2015.12  
A/Map position: 1

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 198;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDDD 443  
DB 101 GDDDDDDDD 108

## RESULT 22

Conserved hypothetical protein AF1309 - Archaeoglobus fulgidus  
C/Species: Archaeoglobus fulgidus  
C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999

C/Accession: D69413  
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A/Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlisch, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Moese, C.R.; Venter, J.C.  
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A/Reference number: A69250; MUID:98049343; PMID:9389475  
A/Accession: D69413  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-211 <KLE>  
A/Cross-references: GB:AE001013; GB:AE000782; NID:g2689336; PIDN:AAB8938.1; PID:g264927

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 211;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 PULLPLAA 336  
DB 174 PULLPLAA 181

## RESULT 23

Antiterminator Q - phage 82  
C/Alternate names: regulatory protein Q

C/Species: phage 82  
C/Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 23-Jul-1999  
C/Accession: A29791; S66585  
R/Goliger, J.A.; Roberts, J.W.  
J. Biol. Chem. 262, 11721-11725, 1987  
A/Title: Bacteriophage 82 gene Q and Q protein: sequence, overproduction, and activity  
A/Reference number: A29791; MUID:87308148; PMID:3624233  
A/Accession: A29791  
A/Molecule type: DNA  
A/Residues: 1-229 <GOL>

A/Cross-references: GB:J02803; NID:g215364; PIDN:AAA32298.1; PID:g215365  
R/Mahdi, A.A.; Sharples, G.J.; Mandal, T.N.; Lloyd, R.G.  
J. Mol. Biol. 257, 561-573, 1996  
A/Title: Holliday junction resolvases encoded by homologous rusa genes in Escherichia co  
A/Reference number: S66579; MUID:96196428; PMID:8648624  
A/Accession: S66585  
A/Molecule type: DNA  
A/Residues: 1-229 <MAH>

A/Cross-references: EMBL:X92588; NID:g1051111; PIDN:CAA63332.1; PID:g1051118  
C/Genetics:  
A/Gene: Q  
C/Superfamily: phage 82 regulatory protein Q

C/Keywords: DNA binding; late protein; transcription regulation

Query Match  
Best Local Similarity 1.8%; Score 8; DB 1; Length 229;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TLERLKKL 297  
DB 142 TLERLKKL 149

## RESULT 24

Antitermination protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0;  
C/Species: Escherichia coli  
C/Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001

C/Accession: D90831  
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
A/Reference number: A99629; MUID:21156231; PMID:11258796  
A/Accession: D90831  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-229 <HAY>  
A/Cross-references: GB:BA000007; PIDN:BA035043.1; PID:g13361084; GSPDB:GN00154  
A/Experimental source: strain O157:H7, substrain RIMD 0509952  
C/Genetics:  
A/Gene: Ecol620  
C/Superfamily: phage 82 regulatory protein Q

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 229;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TLERLKKL 297  
DB 142 TLERLKKL 149

## RESULT 25

G85688

Probable antiterminator Q of prophage CP-933X Z1874 [imported] - Escherichia coli (stra  
C/Species: Escherichia coli  
C/Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C/Accession: G85688  
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; DiMalanta, E.; Potamousis, K.; Apodaca  
Nature 409, 529-533, 2001  
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A/Reference number: A85480; MUID:21074935; PMID:11206551  
A/Accession: G85688  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-229 <STO>

A/Cross-references: GB:AE005174; NID:g12514802; PIDN:AAG55971.1; GSPDB:GN00145; UWGP:Z1  
A/Experimental source: strain O157:H7, substrain EDL933  
C/Genetics:  
A/Gene: Z1874  
C/Superfamily: phage 82 regulatory protein Q

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 229;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TLERLKKL 297  
DB 142 TLERLKKL 149

## RESULT 26

T47271

hypothetical protein tmpr [imported] - Mycobacterium smegmatis insertion sequence IS1096  
C/Species: Mycobacterium smegmatis  
C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 21-Jul-2000  
C/Accession: T47271  
R/Cirillo, J.D.; Barletta, R.G.; Bloom, B.R.; Jacobs, W.R.  
J. Bacteriol. 173, 7772-7780, 1991  
A/Title: A novel transposon trap for mycobacteria: isolation and characterization of IS1  
A/Reference number: Z24439; MUID:92078082; PMID:1660454  
A/Accession: T47271  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-237 <CIR>  
A/Cross-references: EMBL:M76495; NID:g150003; PIDN:AAA98488.1; PID:g150004  
A/Experimental source: insertion sequence IS1096  
C/Genetics:  
A/Gene: tmpr  
A/Mobile element: insertion sequence IS1096

Query Match 1.8%; Score 8; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 431 FDLSEGGD 438  
|||||  
Db 119 FDLSEGGD 126

RESULT 27  
T47611  
hypothetical protein T14E10.90 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 28-Jul-2000  
C/Accession: T47611  
R/Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Lemcke, K.; M  
submitted to the Protein Sequence Database, March 2000  
A/Reference number: Z24463  
A/Accession: T47611  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-250 <OBE>  
A/Cross-references: EMBL:AL138656  
A/Experimental source: cultivar Columbia; BAC clone T14E10  
C/Genetics:  
A/Map position: 3  
A/Introns: 73/3; 129/3; 188/3  
A/Note: T14E10.90  
C/Superfamily: Arabidopsis thaliana hypothetical protein T14E10.90

Query Match 1.8%; Score 8; DB 2; Length 250;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 RLKLERD 300  
|||||  
Db 54 RLKLERD 61

RESULT 28  
G84645  
hypothetical protein At2g25220 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C/Accession: G84645  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Mofatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: G84645  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-260 <STO>

A/Cross-references: GB:AE002093; NID:g4567255; PIDN:AAD23669.1; GSPDB:GN00139  
C/Genetics:  
A/Gene: At2g25220  
A/Map position: 2

Query Match 1.8%; Score 8; DB 2; Length 260;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRLKS 136  
|||||  
Db 125 VIHRLKS 132

RESULT 29  
A34599  
DNA-binding protein MCM1 - yeast (Saccharomyces cerevisiae)  
N/Alternate names: DNA-binding protein FUN80; DNA-binding protein PRTF/GRM; protein YW9;  
C/Species: Saccharomyces cerevisiae  
C/Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 31-Mar-2000  
C/Accession: A34599; S14888; S22855; S52892  
R/Ammeyer, G.  
Genes Dev. 4, 299-312, 1990  
A/Title: Identification, purification, and cloning of a polypeptide (PRTF/GRM) that bin  
A/Reference number: A34599; MUID:90249735; PMID:2159934  
A/Accession: A34599  
A/Molecule type: DNA  
A/Residues: 1-286 <AMM>  
A/Cross-references: GB:X52453; NID:g5266; PIDN:CAA36691.1; PID:g5267  
R/Passmore, S.; Maine, G.T.; Bible, R.; Christ, C.; Tye, B.K.  
J. Mol. Biol. 204, 593-606, 1988  
A/Title: Saccharomyces cerevisiae protein involved in plasmid maintenance is necessary  
A/Reference number: S14888; MUID:89141759; PMID:3066908  
A/Accession: S14888  
A/Molecule type: DNA  
A/Residues: 1-286 <PAS>  
A/Cross-references: EMBL:X14187; NID:g3909; PIDN:CAA32389.1; PID:g3910  
R/Dubois, E.; Bercy, J.; Descamps, F.; Messenguy, F.  
Gene 55, 265-275, 1987  
A/Title: Characterization of two new genes essential for vegetative growth in Saccharom  
A/Reference number: S22855; MUID:88030692; PMID:3311883  
A/Accession: S22855  
A/Molecule type: DNA  
A/Residues: 1-8,'S',10-36,'F',38-155,'AR' <DUB>  
A/Cross-references: EMBL:M17511; NID:g171524; PIDN:AAA34609.1; PID:g171525  
A/Note: the authors translated the codon TCT for residue 9 as Pro and TTC for residue 3  
R/Odell, C.; Bowman, S.  
submitted to the EMBL Data Library, February 1995  
A/Reference number: S52885  
A/Accession: S52892  
A/Molecule type: DNA  
A/Residues: 1-286 <ODE>  
A/Cross-references: EMBL:Z48502; NID:g695715; PIDN:CAA88409.1; PID:g695723; MIPS:YMR043  
C/Genetics:  
A/Gene: SGD:MCM1  
A/Cross-references: SGD:S0004646; MIPS:YMR043w  
A/Map position: 13R  
C/Superfamily: transcription factor squa; serum response factor DNA-binding domain homo  
C/Keywords: DNA binding; nucleus; transcription regulation  
F/17-72/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match 1.8%; Score 8; DB 2; Length 286;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDDD 443  
|||||  
Db 108 GDDDDDDDD 115

RESULT 30  
A37818  
osteopontin precursor - mouse

N;Alternate names: bone sialoprotein I; early T lymphocyte activation 1 protein; phospho  
C;Species: Mus musculus (house mouse)  
C;Date: 12-Feb-1993 #sequence\_revision 27-Jun-1994 #text\_change 10-Sep-1999  
C;Accession: A37818; S04078; S12064; A33853; J10105; A60331; S11677  
R;Miyazaki, Y.; Setoguchi, M.; Yoshida, S.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.  
J. Biol. Chem. 265, 14432-14438, 1990  
A;Title: The mouse osteopontin gene. Expression in monocytic lineages and complete nucle  
A;Reference number: A37818; MUID:90354433; PMID:2387863  
A;Accession: A37818  
A;Molecule type: DNA  
A;Residues: 1-294 <MT>  
A;Cross-references: EMBL:X51834  
R;Miyazaki, Y.; Setoguchi, M.; Yoshida, S.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.  
Nucleic Acids Res. 17, 3298, 1989  
A;Title: Nucleotide sequence of cDNA for mouse osteopontin-like protein.  
A;Reference number: S04078; MUID:89263742; PMID:2726465  
A;Accession: S04078  
A;Molecule type: mRNA  
A;Residues: 1-294 <MT>  
A;Cross-references: EMBL:X13986; NID:G53755; PIDN:CAA32165.1; PID:G53756  
Yamamoto, S.  
Submitted to the EMBL Data Library, January 1990  
A;Reference number: S12064  
A;Accession: S12064  
A;Molecule type: DNA  
A;Residues: 1-121, 'F', 123-294 <YAM>  
A;Cross-references: EMBL:X51834; NID:G53520; PIDN:CAA36132.1; PID:G297546  
R;Craig, A.M.; Smith, J.H.; Denhardt, D.T.  
J. Biol. Chem. 264, 9682-9689, 1989  
A;Title: Osteopontin, a transformation-associated cell adhesion phosphoprotein, is induc  
A;Reference number: A33853; MUID:89255479; PMID:2722855  
A;Accession: A33853  
A;Molecule type: mRNA  
A;Residues: 1-98, 'G', 100-294 <CRA>  
A;Cross-references: GB:J04806; NID:G200157; PIDN:AA57265.1; PID:G200158  
R;Patarca, R.; Freeman, G.J.; Singh, R.P.; Wei, F.Y.; Duffee, T.; Blatner, F.; Regnier, J.  
Exp. Med. 170, 145-161, 1989  
A;Title: Structural and functional studies of the early T lymphocyte activation 1 (Eta-1  
on.  
A;Reference number: J10105; MUID:89310352; PMID:2787378  
A;Accession: J10105  
A;Molecule type: mRNA  
A;Residues: 1-42, 'P', 44-294 <PAT>  
A;Cross-references: EMBL:X16151; NID:G50863; PIDN:CAA34276.1; PID:G50864  
R;Singh, R.P.; Patarca, R.; Schwartz, J.; Singh, P.; Cantor, H.  
J. Exp. Med. 171, 1931-1942, 1990  
A;Title: Definition of a specific interaction between the early T lymphocyte activation  
A;Reference number: A60931; MUID:90278349; PMID:2351930  
A;Accession: A60931  
A;Molecule type: protein  
A;Residues: 158-176 <SIN>  
C;Comment: This protein is an acidic glycoprotein rich in aspartic acid, glutamic acid,  
C;Genetics:  
A;Gene: Eta-1  
A;Map position: 5  
A;Introns: 18/3; 30/3; 57/3; 71/3; 165/3  
C;Superfamily: osteopontin  
C;Keywords: bone; cell binding; extracellular matrix; phosphoprotein; sialoglycoprotein  
F;1-16/Domain: signal sequence #status predicted <SIG>  
F;17-294/Product: osteopontin #status predicted <SIG>  
F;85-96/Region: aspartic acid-rich  
F;144-146/Region: cell attachment (R-G-D) motif  
F;78/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 1.8%; Score 8; DB 1; Length 294;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
DB 88 DDDDDDDG 95

RESULT 31  
T52117  
zinc finger protein [imported] - Arabidopsis thaliana (fragment)  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000  
C;Accession: T52117  
R;Bilodeau, P.  
Submitted to the EMBL Data Library, October 1997  
A;Reference number: Z25963  
A;Accession: T52117  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-298 <BIL>  
A;Cross-references: EMBL:AF030304; PIDN:AAC09174.1

Query Match 1.8%; Score 8; DB 2; Length 298;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
DB 25 DDDDDDDG 32

RESULT 32  
JC5811  
osteopontin - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 04-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 07-May-1999  
C;Accession: JC5811  
R;Iasa, M.; Chang, P.L.; Prince, C.W.; Pina, L.A.  
Biochem. Biophys. Res. Commun. 240, 602-605, 1997  
A;Title: Phosphorylation of osteopontin by Golgi apparatus casein kinase.  
A;Reference number: JC5811; MUID:98063283; PMID:9398611  
A;Accession: JC5811  
A;Molecule type: protein  
A;Residues: 1-301 <IAS>  
A;Experimental source: brain  
C;Comment: This protein is involved in the initiation of the bone calcification process,  
ues, early resistance to bacterial infection and binding of tumor cells at secondary sit  
C;Superfamily: osteopontin

Query Match 1.8%; Score 8; DB 2; Length 301;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
DB 72 DDDDDDDG 79

RESULT 33  
A25917  
osteopontin precursor - rat  
N;Alternate names: bone sialoprotein I; phosphoprotein I, secreted; tumor-secreted phospho  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 18-Dec-1987 #sequence\_revision 27-Jun-1994 #text\_change 10-Sep-1999  
C;Accession: A25917; A45132; B45132; S28772; S04506; A45925  
R;Oldberg, A.; Franzen, A.; Heinegard, D.  
Proc. Natl. Acad. Sci. U.S.A. 83, 8819-8823, 1986  
A;Title: Cloning and sequence analysis of rat bone sialoprotein (osteopontin) cDNA reve  
A;Reference number: A25917; MUID:87067405; PMID:3024151  
A;Accession: A25917  
A;Molecule type: mRNA  
A;Residues: 1-317 <OLD>  
A;Cross-references: GB:M14656; NID:G205859; PIDN:AAA41762.1; PID:G205860  
R;Singh, K.; Mukherjee, A.B.; De Vouge, M.W.; Mukherjee, B.B.  
J. Biol. Chem. 267, 23847-23851, 1992  
A;Title: Differential processing of osteopontin transcripts in rat kidney- and osteobla  
A;Reference number: A45132; MUID:93054745; PMID:1429723  
A;Accession: A45132  
A;Molecule type: protein  
A;Residues: 36-51 <SIN1>



A/Experimental source: kidney  
A/Note: sequence extracted from NCBI backbone (NCBIP:118869)  
A/Accession: B45132  
A/Molecule type: protein  
A/Residues: 272-282 <SIN2>  
A/Note: sequence extracted from NCBI backbone (NCBIP:118871)  
R/Prince, C.W.; Oosawa, T.; Butler, W.T.; Tomana, M.; Bhowm, A.S.; Bhowm, M.; Schrehenld  
J. Biol. Chem. 262, 2900-2907, 1987  
A/Title: Isolation, characterization, and biosynthesis of a phosphorylated glycoprotein  
A/Reference number: S28772; MUID:87137549; PMID:3469201  
A/Accession: S28772  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 17-26, 'X' <PRI>  
R/Senger, D.R.; Petruzzi, C.A.; Papadopoulos, A.; Tenen, D.G.  
Biochim. Biophys. Acta 996, 43-48, 1989  
A/Title: Purification of a human milk protein closely similar to tumor-secreted phosphog  
A/Reference number: S04505; MUID:89287357; PMID:2736258  
A/Accession: S04506  
A/Status: preliminary  
A/Molecule type: protein  
Residues: 17-26, 155-167 <SEN>  
R/Senger, D.R.; Petruzzi, C.A.; Gracey, C.F.; Papadopoulos, A.; Tenen, D.G.  
Cancer Res. 48, 5770-5774, 1988  
A/Title: Secreted phosphoproteins associated with neoplastic transformation: close homol  
A/Reference number: A45925; MUID:89002730; PMID:3167835  
A/Accession: A45925  
A/Molecule type: protein  
A/Residues: 17-25 <SE2>  
C/Comment: This protein is an acidic glycoprotein rich in aspartic acid, glutamic acid,  
C/Superfamily: osteopontin  
C/Keywords: bone; cell binding; extracellular matrix; phosphoprotein; sialoglycoprotein  
F/1-16/Domain: signal sequence #status predicted <SIG>  
F/17-317/Product: osteopontin #status predicted <MAT>  
F/86-96/Region: aspartic acid-rich  
F/144-146/Region: cell attachment (R-G-D) motif  
F/79/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.8%; Score 8; DB 1; Length 317;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
Db 88 DDDDDDDG 95

RESULT 34  
9026  
Observed hypothetical protein MTH1196 - Methanobacterium thermoautotrophicum (strain De  
Species: Methanobacterium thermoautotrophicum  
C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 26-May-2000  
C/Accession: E69026  
R/Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
/Qu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.  
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funcn  
A/Reference number: A69000; MUID:98037514; PMID:9371463  
A/Accession: E69026  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-318 <MTH>  
A/Cross-references: GB:AE000888; GB:AE000666; NID:g2622304; PIDN:AAB85685.1; PID:g262230  
A/Experimental source: strain Delta H  
C/Genetics:  
A/Gene: MTH1196  
A/Start codon: TTG  
C/Superfamily: conserved hypothetical protein MTH1196

Query Match 1.8%; Score 8; DB 2; Length 318;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 VKKLKIE 51  
Db 28 VKKLKIE 35

## RESULT 35

hypothetical protein R13F6.7 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Mar-2000  
C/Accession: T16747  
R/Miller, N.  
submitted to the EMBL Data Library, April 1994  
A/Description: The sequence of C. elegans cosmid R13F6.  
A/Reference number: Z18570  
A/Accession: T16747  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-328 <ML>  
A/Cross-references: EMBL:U00046; NID:g470358; PID:g470364; PIDN:AAC47047.1; GSPDB:GN000.  
A/Experimental source: strain Bristol N2; clone R13F6  
C/Genetics:  
A/Gene: CESP.R13F6.7  
A/Map position: 3  
A/Introns: 20/3; 160/3; 222/2; 286/2  
C/Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 i

Query Match 1.8%; Score 8; DB 2; Length 328;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAEILS 58  
Db 90 EKEAEILS 97

## RESULT 36

R5HSL0  
ribosomal protein L10 [similarity] - Halobacterium salinarum  
N/Alternate names: ribosomal protein P0  
C/Species: Halobacterium salinarum  
C/Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 21-Jul-2000  
C/Accession: S01315; S04120; S11587  
R/Itoh, T.  
Eur. J. Biochem. 176, 297-303, 1988  
A/Title: Complete nucleotide sequence of the ribosomal 'A' protein operon from the arch.  
A/Reference number: S01314; MUID:88329082; PMID:2458258  
A/Accession: S01315  
A/Molecule type: DNA  
A/Residues: 1-352 <ITO>  
A/Cross-references: EMBL:X13008; NID:g43532; PIDN:CAA31431.1; PID:g43534  
A/Experimental source: strain S9  
A/Note: the source is designated as Halobacterium halobium  
R/Shimmin, L.C.; Dennis, P.P.  
EMBO J. 8, 1225-1235, 1989  
A/Title: Characterization of the L11, L1, L10 and L12 equivalent ribosomal protein gene  
A/Reference number: S04116; MUID:89305527; PMID:2743981  
A/Accession: S04120  
A/Molecule type: DNA  
A/Residues: 1-58, 'V', 60-352 <SHI>  
A/Cross-references: EMBL:X15078; NID:g43449; PIDN:CAA33180.1; PID:g43454  
A/Note: the source is designated as Halobacterium cutirubrum  
C/Superfamily: rat acidic ribosomal protein P0  
C/Keywords: protein biosynthesis; ribosome

Query Match 1.8%; Score 8; DB 1; Length 352;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
Db 333 DDDDDDDG 340

## RESULT 37

G84266

50S ribosomal protein L10P [imported] - Halobacterium sp. NRC-1

C/Species: Halobacterium sp. NRC-1

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001

C/Accession: G84266

R/Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li

A/Title: Genome sequence of Halobacterium species NRC-1.

A/Reference number: A84160; MUID:20504483; PMID:11016950

A/Accession: G84266

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-352 &lt;STO&gt;

A/Cross-references: GB:AE004437; NID:g10580651; PIDN:AAG19499.1; GSPDB:GN00138

C/Genetics:

A/Gene: rpl10p

C/Superfamily: rat acidic ribosomal protein P0

## Query Match

Best Local Similarity 1.8%; Score 8; DB 2; Length 352;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444

Db 333 DDDDDDDG 340

## RESULT 38

G96668

protein F1N19.7 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 01-Mar-2002

C/Accession: G96668

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizart, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

R.; M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: G96668

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-368 &lt;STO&gt;

A/Cross-references: GB:AE005173; NID:g6633811; PIDN:AAF19670.1; GSPDB:GN00141

C/Genetics:

A/Gene: F1N19.7

A/Map position: 1

C/Superfamily: Arabidopsis thaliana hypothetical protein T4F9.90

## Query Match

Best Local Similarity 1.8%; Score 8; DB 2; Length 368;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444

Db 61 DDDDDDDG 68

## RESULT 39

T33173

hypothetical protein C24G6.5 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C/Accession: T33173

R/Greco, T.; Bradshaw, H.; Keppler, D.

submitted to the EMBL Data Library, May 1998

A/Description: The sequence of C. elegans cosmid C24G6.

A/Reference number: 221298

A/Accession: T33173

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-374 &lt;GRE&gt;

A/Cross-references: EMBL:AF067936; PIDN:AACT19208.1; GSPDB:GN00023; CESP:C24G6.5

A/Experimental source: strain Bristol N2; clone C24G6

C/Genetics:

A/Gene: CESP:C24G6.5

A/Map position: 5

A/Introns: 83/1; 351/3

## Query Match

Best Local Similarity 1.8%; Score 8; DB 2; Length 374;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDD 443

Db 189 GDDDDDDD 196

## RESULT 40

T04645

hypothetical protein F10N7.190 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 20-Sep-1999

C/Accession: T04645

R/Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.

submitted to the Protein Sequence Database, March 1999

A/Reference number: 215263

A/Accession: T04645

A/Molecule type: DNA

A/Residues: 1-379 &lt;BEV&gt;

A/Cross-references: EMBL:AL021636

A/Experimental source: cultivar Columbia; BAC clone F10N7

C/Genetics:

A/Map position: 4

A/Introns: 48/1; 221/1; 252/3; 287/1; 326/3

A/Note: F10N7.190

C/Superfamily: kinase-related transforming protein; protein kinase homology

## Query Match

Best Local Similarity 1.8%; Score 8; DB 2; Length 379;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VTHRDLLK 136

Db 233 VTHRDLLK 240

## RESULT 41

T01451

protein kinase homolog F24O1.13 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001

C/Accession: T01451

R/Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Co

eologis, A.; Ecker, J.R.

submitted to the EMBL Data Library, January 1998

A/Description: Genomic sequence for Arabidopsis thaliana BAC F24O1.

A/Reference number: 214211

A/Accession: T01451

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-390 &lt;SHI&gt;

A/Cross-references: EMBL:AC003113; NID:g2689438; PID:g2781357; GSPDB:GN00059; ATSP:F24O

C/Genetics:

A/Gene: ATSP:F24O1.13

A;Map position: 1  
A;Introns: 149/3; 301/3  
C;Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 1.8%; Score 8; DB 2; Length 390;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VTHRDLS 136  
|||||  
Db 208 VTHRDLS 215

## RESULT 42

D64366  
hypothetical protein homolog MJ0532 - Methanococcus jannaschii

C;Species: Methanococcus jannaschii  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C;Accession: D64366

R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glodek, A.; Bon, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996

A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A;Reference number: A64300; MUID:96337999; PMID:8688087

A;Accession: D64366  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-391 <BUL>

C;Cross-references: GB:U67502; GB:L77117; NID:g2826293; PIDN:AAB98523.1; PID:g1591235; T

A;Map position: REV468785-467610  
C;Superfamily: hypothetical protein MJ0532

Query Match 1.8%; Score 8; DB 2; Length 391;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVIADGV 146  
|||||  
Db 149 VVIADGV 156

## RESULT 43

S52578  
serine/threonine-specific protein kinase NPK15 (EC 2.7.1.-) - common tobacco

C;Species: Nicotiana tabacum (common tobacco)  
C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 24-Sep-1999

A;Accession: S52578

Mito, Y.; Banno, H.; Moribe, T.; Hinata, K.; Machida, K.H.Y.  
Mol. Gen. Genet. 245, 1-10, 1994

A;Title: NPK15, a tobacco protein-serine/threonine kinase with a single hydrophobic regi

A;Reference number: S52578; MUID:95147840; PMID:7845351

A;Accession: S52578  
A;Molecule type: mRNA

A;Residues: 1-422 <ITO>  
A;Cross-references: EMBL:D31737; NID:g505145; PIDN:BA06538.1; PID:d1007109; PID:g505146

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F;114-382/Domain: protein kinase homology <KIN>

F;122-130/Region: protein kinase ATP-binding motif

QY 129 VTHRDLS 136  
|||||  
Db 235 VTHRDLS 242

## RESULT 44

T05676  
hypothetical protein F20M13.40 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 23-Jul-1999

C;Accession: T05676

R;Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.I

submitted to the Protein Sequence Database, February 1999

A;Reference number: Z15420

A;Accession: T05676

A;Molecule type: DNA

A;Residues: 1-426 <BEV>

A;Cross-references: EMBL:AL035540  
A;Experimental source: cultivar Columbia; BAC clone F20M13

C;Genetics:

A;Map position: 4

A;Introns: 90/3; 149/2; 204/1; 350/3  
A;Note: F20M13.40

QY 436 GDDDDDD 443  
|||||  
Db 407 GDDDDDD 414

## RESULT 45

B54843  
nemo, form II - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 19-Dec-1997

C;Accession: B54843

R;Choi, K.W.; Benzer, S.  
Cell 78, 125-136, 1994

A;Title: Rotation of photoreceptor clusters in the developing Drosophila eye requires ti

A;Reference number: A54843; MUID:94306509; PMID:8033204

A;Accession: B54843  
A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-434 <CHO>

A;Cross-references: GB:U12009

C;Genetics:

A;Gene: nmo

A;Cross-references: FlyBase:FBgn0011817  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

C;Keywords: alternative splicing; ATP

F;38-301/Domain: protein kinase homology <KIN>

F;46-54/Region: protein kinase ATP-binding motif

Query Match 1.8%; Score 8; DB 2; Length 434;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
|||||  
Db 179 VLKICDFG 186

## RESULT 46

T51402  
serine/threonine-specific protein kinase-like protein - Arabidopsis thaliana

N;Alternate names: protein F14F8\_110  
C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 17-Nov-2000

C;Accession: T51402

R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Me

submitted to the Protein Sequence Database, August 2000

A;Reference number: Z25394

A;Accession: T51402

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-436 <SAT>

A/Cross-references: EMBL:AL391144  
A/Experimental source: cultivar Columbia; BAC clone F14F8  
C/Genetics:  
A/Map position: 5  
A/Introns: 107/2; 158/3; 202/1; 254/3; 322/3  
A/Note: F14F8.110  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

Query Match 1.8%; Score 8; DB 2; Length 436;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 129 VHRDLKS 136  
Db 235 VHRDLKS 242

## RESULT 47

T31734  
Hypothetical protein T05C3.5 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 31-Jan-2000  
C/Accession: T31734  
R/Blanchard, M.; Bradshaw, H.; Kramer, J.  
submitted to the EMBL Data Library, July 1997  
A/Description: The sequence of C. elegans cosmid T05C3.  
A/Reference number: 221076  
A/Accession: T31734  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-439 <BLA>  
A/Cross-references: EMBL:AF016428; PIDN:AAB5361.1; GSPDB:GN00023; CESP:T05C3.5  
A/Experimental source: strain Bristol N2; clone T05C3  
C/Genetics:  
A/Gene: CESP:T05C3.5  
A/Map position: 5  
A/Introns: 267/3; 298/2; 377/3  
C/Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

Query Match 1.8%; Score 8; DB 2; Length 439;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDD 443  
Db 105 GDDDDDD 112

## RESULT 48

T39232  
Probable serine threonine protein kinase - fission yeast (Schizosaccharomyces pombe)  
C/Species: Schizosaccharomyces pombe  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C/Accession: T39232  
R/Churcher, C.M.; Gentles, S.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, August 1997  
A/Reference number: 221837  
A/Accession: T39232  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-471 <CHU>  
A/Cross-references: EMBL:Z98763; PIDN:CAB11493.1; GSPDB:GN00066; SPDB:SPAC9G1.09  
A/Experimental source: strain 972h-; cosmid c9G1  
C/Genetics:  
A/Gene: SPDB:SPAC9G1.09  
A/Map position: 1  
A/Introns: 39/3  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

Query Match 1.8%; Score 8; DB 2; Length 471;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
Db 171 WMAPEVIQ 178

## RESULT 49

A54843  
nemo, form I - fruit fly (Drosophila melanogaster)  
C/Species: Drosophila melanogaster  
C/Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 24-Sep-1999  
C/Accession: A54843  
R/Choi, K.W.; Benzer, S.  
Cell 78, 125-136, 1994  
A/Title: Rotation of photoreceptor clusters in the developing Drosophila eye requires t  
A/Reference number: A54843; MUID:94306509; PMID:8033204  
A/Accession: A54843  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-477 <CHO>  
A/Cross-references: GB:U12009; NID:G515669; PIDN:AAA21124.1; PID:G532558  
C/Genetics:  
A/Gene: nmo  
A/Cross-references: FlyBase:Fgn0011817  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol  
C/Keywords: alternative splicing; ATP  
F;46-54/Region: protein kinase ATP-binding motif

Query Match 1.8%; Score 8; DB 2; Length 477;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 146 VLKICDFG 153  
Db 179 VLKICDFG 186

## RESULT 50

S37845  
transcription initiation factor IIE chain TPA1 - Yeast (Saccharomyces cerevisiae)  
N/Alternate names: protein YKL028w; transcription initiation factor a 66k chain  
C/Species: Saccharomyces cerevisiae  
A/Variety: strain S288C  
C/Date: 23-Feb-1995 #sequence\_revision 11-Aug-1995 #text\_change 21-Jul-2000  
C/Accession: S37845; A55465; C55465  
R/Rieger, M.  
submitted to the Protein Sequence Database, March 1994  
A/Reference number: S37832  
A/Accession: S37845  
A/Molecule type: DNA  
A/Residues: 1-482 <RIE>  
A/Cross-references: EMBL:Z28028; NID:G486026; PID:G486027; MIPS:YKL028w  
R/Feaver, W.J.; Henry, N.L.; Bushnell, D.A.; Sayre, M.H.; Brickner, J.H.; Gileadi, O.;  
J. Biol. Chem. 269, 27549-27553, 1994  
A/Title: Yeast TRIE. Cloning, expression, and homology to vertebrate proteins.  
A/Reference number: A55465; MUID:95050500; PMID:7961670  
A/Accession: A55465  
A/Molecule type: DNA  
A/Residues: 1-482 <FEA>  
A/Cross-references: GB:U12825; NID:G607957; PIDN:AAA62665.1; PID:G607958  
A/Accession: C55465  
A/Molecule type: protein  
A/Residues: 102-108; 206-220 <FE2>  
C/Genetics:  
A/Gene: SGD:TPA1  
A/Cross-references: SGD:S0001511; MIPS:YKL028w  
A/Map position: 11L  
C/Keywords: nucleus; transcription initiation

Query Match 1.8%; Score 8; DB 2; Length 482;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



OY 436 GDDDDDDDD 443  
|||||  
Db 464 GDDDDDDDD 471

## RESULT 51

T49237  
hypothetical protein F7K15.80 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C/Accession: T49237  
R/Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Rudd, S.; Ler  
submitted to the Protein Sequence Database, April 2000  
A/Reference number: Z25019  
A/Accession: T49237  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-485 <OBE>  
A/Cross-references: EMBL:AL353871; GSPDB:GN00061; ATSP:F7K15.80  
A/Experimental source: cultivar Columbia; BAC clone F7K15  
A/Genetics:  
A/Gene: ATSP:F7K15.80  
A/Map position: 3  
A/Introns: 292/3; 359/3; 425/1

Query Match 1.8%; Score 8; DB 2; Length 485;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
|||||  
Db 25 DDDDDDDG 32

## RESULT 52

D84860  
probable protein kinase [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C/Accession: D84860  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
eus, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MUID:20063487; PMID:10617197  
A/Accession: D84860  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-494 <STO>  
A/Cross-references: GB:AE002093; NID:G4512659; PIDN:AAD21713.1; GSPDB:GN00139  
A/Genetics:  
A/Gene: At2g42960  
A/Map position: 2

Query Match 1.8%; Score 8; DB 2; Length 494;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 EVAVKKL 48  
|||||  
Db 207 EVAVKKL 214

## RESULT 53

T10558  
hypothetical protein T12G13.150 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 15-Oct-1999  
C/Accession: T10558  
R/Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro  
submitted to the Protein Sequence Database, June 1999  
A/Reference number: Z16533

A/Accession: T10558  
A/Molecule type: DNA  
A/Residues: 1-504 <BEV>  
A/Cross-references: EMBL:AL080252; GSPDB:GN00062; ATSP:T12G13.150  
A/Experimental source: cultivar Columbia; BAC clone T12G13  
A/Genetics:  
A/Gene: ATSP:T12G13.150  
A/Map position: 4  
A/Introns: 54/2; 89/3; 192/2; 227/3; 334/3; 360/2; 402/1; 496/1

Query Match 1.8%; Score 8; DB 2; Length 504;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 DKEVAVKK 46  
|||||  
Db 274 DKEVAVKK 281

## RESULT 54

B96524  
hypothetical protein F11A17.5 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
C/Accession: B96524  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: B96524  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-513 <STO>  
A/Cross-references: GB:AE005173; NID:G5733869; PIDN:AAD49757.1; GSPDB:GN00141  
A/Genetics:  
A/Gene: F11A17.5  
A/Map position: 1  
C/Superfamily: Arabidopsis thaliana hypothetical protein F17J16.30

Query Match 1.8%; Score 8; DB 2; Length 513;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
|||||  
Db 319 DDDDDDDG 326

## RESULT 55

T47786  
hypothetical protein F17J16.90 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C/Accession: T47786  
R/D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzini, M.; Valle, G.; Mewes, H.W.; Rudd, S.;  
submitted to the Protein Sequence Database, April 2000  
A/Reference number: Z24476  
A/Accession: T47786  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-526 <DAN>  
A/Cross-references: EMBL:AL163527  
A/Experimental source: cultivar Columbia; BAC clone F17J16  
A/Genetics:  
A/Map position: 3  
A/Introns: 67/3; 247/3; 289/2; 322/2

A>Note: F17J16.90

Query Match 1.8%; Score 8; DB 2; Length 526;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDDEDDDD 442  
|||||  
Db 491 EGDDEDDDD 498

## RESULT 56

C83284

Probable biotin-dependent carboxylase PA2888 [imported] - Pseudomonas aeruginosa (strain C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C/Accession: C83284

R/Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coultet, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim, Lory, S.; Olson, M.V.  
ature 406, 959-964, 2000

A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A/Reference number: A82950; MUID:20437337; PMID:10984043

A/Accession: C83284

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-538 &lt;STO&gt;

A/Cross-references: GB:AE004715; GB:AE004091; NID:99948977; PIDN:AAG06276.1; GSPDB:GN001  
A/Experimental source: strain PA01

C/Genetics:

A/Gene: PA2888

C/Superfamily: propionyl-CoA carboxylase beta chain

Query Match 1.8%; Score 8; DB 2; Length 538;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 324 EQSNTPL 331  
|||||  
Db 364 EQSNTPL 371

## RESULT 57

A12238

glutamine-fructose-6-P-aminotransferase [imported] - Nostoc sp. (strain PCC 7120)  
C/Species: Nostoc sp.

A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
A/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
A/Accession: A12238

A/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001

A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A/Reference number: AB1807; MUID:21595285; PMID:11759840

A/Accession: A12238

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-541 &lt;KUR&gt;

A/Cross-references: GB:BA000019; PIDN:BA075163.1; PID:gl7132597; GSPDB:GN00179  
A/Experimental source: strain PCC 7120

C/Genetics:

A/Gene: nodM

C/Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)

Query Match 1.8%; Score 8; DB 2; Length 541;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 EIEATLER 293  
|||||  
Db 375 EIEATLER 382

## RESULT 58

T49839

hypothetical protein B24H17.250 [imported] - Neurospora crassa  
C/Species: Neurospora crassa

C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C/Accession: T49839

R/Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura  
submitted to the Protein Sequence Database, May 2000

A/Reference number: Z25022

A/Accession: T49839

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-560 &lt;SCH&gt;

A/Cross-references: EMBL:AL356815; GSPDB:GN00116; NCSP:B24H17.250  
A/Experimental source: BAC clone B24H17; strain OR74A

C/Genetics:

A/Gene: NCSP:B24H17.250

A/Map position: 6

A/Introns: 72/1; 142/3; 270/1; 310/1; 380/3; 411/2; 446/2; 468/1; 485/3

Query Match 1.8%; Score 8; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
|||||  
Db 385 DDDDDDDG 392

## RESULT 59

JC5957

transforming growth factor-beta activated kinase (EC 2.7.-.-) 1c - human  
C/Species: Homo sapiens (man)

C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 17-Mar-2000  
C/Accession: JC5957

R/Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.  
Biochem. Biophys. Res. Commun. 243, 545-549, 1998

A/Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-in  
A/Reference number: JC5955; MUID:98153801; PMID:9480845

A/Accession: JC5957

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-567 &lt;SAK&gt;

A/Cross-references: DDBJ:AB009358  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

C/Keywords: phosphotransferase

Query Match 1.8%; Score 8; DB 2; Length 567;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
|||||  
Db 170 VLKICDFG 177

## RESULT 60

JC5955

transforming growth factor-beta activated kinase (EC 2.7.-.-) 1a - human  
C/Species: Homo sapiens (man)

C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C/Accession: JC5955

R/Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.  
Biochem. Biophys. Res. Commun. 243, 545-549, 1998

A/Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-in  
A/Reference number: JC5955; MUID:98153801; PMID:9480845

A/Accession: JC5955

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-579 &lt;SAK&gt;

A/Cross-references: DDBJ:AB009356; NID:g2924623; PIDN:BAA25025.1; PID:g2924624  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol  
C/Keywords: phosphotransferase

Query Match 1.8%; Score 8; DB 2; Length 579;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 146 VLKICDFG 153  
|||||  
Db 170 VLKICDFG 177

## RESULT 61

T49206  
phosphoinositide-specific phospholipase C-like protein - Arabidopsis thaliana  
N/Alternate names: protein F27K19.120  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 28-Jul-2000  
C/Accession: T49206  
R/Benes, V.; Wurmback, E.; Drzonek, H.; Ansoerge, W.; Mewes, H.W.; Rudd, S.; Lemcke, K.;  
submitted to the Protein Sequence Database, April 2000  
1;Reference number: Z25014  
A/Accession: T49206  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-584 <BEN>  
A/Cross-references: EMBL:AL163832; GSPDB:GN00061; ATSP:F27K19.120  
A/Experimental source: cultivar Columbia; BAC clone F27K19  
C/Genetics:  
A/Gene: ATSP:F27K19.120  
A/Map position: 3  
A/Intons: 100/3; 166/2; 211/3; 290/3; 371/2; 410/3; 458/3; 487/3  
C/Superfamily: Arabidopsis thaliana 1-phosphatidylinositol-4,5-bisphosphate phosphodiester  
bisphosphate phosphodiesterase domain Y homology  
F;105-248/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain X hom

Query Match 1.8%; Score 8; DB 2; Length 584;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443  
|||||  
Db 295 GDDDDDD 302

## RESULT 62

S57594  
hypothetical protein YMR227C - yeast (Saccharomyces cerevisiae)  
N/Alternate names: hypothetical protein YMR227C.09c  
C/Species: Saccharomyces cerevisiae  
C/Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 29-Oct-1995  
A/Accession: S57594  
R/Skelton, J.; Church, C.M.  
submitted to the EMBL Data Library, June 1995  
A/Reference number: S57587  
A/Accession: S57594  
A/Molecule type: DNA  
A/Residues: 1-590 <SKE>  
A/Cross-references: EMBL:Z49939; NID:g887599; PIDN:CAA90198.1; PID:g887608; MIPS:YMR227C  
A/Experimental source: strain AB972  
C/Genetics:  
A/Gene: SGD:TAF67  
A/Cross-references: SGD:S0004840; MIPS:YMR227C  
A/Map position: 13R

Query Match 1.8%; Score 8; DB 2; Length 590;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
|||||  
Db 413 DDDDDDDG 420

## RESULT 63

## JC5956

transforming growth factor-beta activated kinase (EC 2.7.-.-) 1b - human  
C/Species: Homo sapiens (man)  
C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C/Accession: JC5956  
R/Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.  
Biochem. Biophys. Res. Commun. 243, 545-549, 1998  
A/Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-in  
A/Reference number: JC5955; MUID:98153801; PMID:9480845  
A/Accession: JC5956  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-606 <SAK>  
A/Cross-references: DDBJ:AB009357; NID:g2924625; PIDN:BAA25026.1; PID:g2924626  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol  
C/Keywords: phosphotransferase

Query Match 1.8%; Score 8; DB 2; Length 606;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 146 VLKICDFG 153  
|||||  
Db 170 VLKICDFG 177

## RESULT 64

C96596  
hypothetical protein T1813.3 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: C96596  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: C96596  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-684 <STO>  
A/Cross-references: GB:AE005173; NID:g11094789; PIDN:AAG29721.1; GSPDB:GN00141  
C/Genetics:  
A/Gene: T1813.3  
A/Map position: 1

Query Match 1.8%; Score 8; DB 2; Length 684;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443  
|||||  
Db 162 GDDDDDD 169

## RESULT 65

C96572  
Protein F12M16.4 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: C96572  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: C96572  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-690 <STO>  
A:Cross-references: GB:AE005173; NID:g7769851; PIDN:AAF69529.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F12M16.4  
A:Map position: 1

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 690;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 172 WMAPEVIO 179  
Db 187 WMAPEVIO 194

RESULT 66  
G02838  
enhancer-of-zeste homolog 2 - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 17-Jul-1998  
C:Accession: G02838  
R:Antonarakis, S.E.  
submitted to the EMBL Data Library, February 1996  
A:Reference number: H01746  
A:Accession: G02838  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-746 <ANT>  
A:Cross-references: EMBL:X95653; NID:g1438063  
C:Genetics:  
A:Gene: GDB:EZH2; EZH1  
A:Cross-references: GDB:701613; OMIM:601573  
A:Map position: 21q22.2-21q22.2

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 746;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDDG 444  
Db 183 DDDDDDDG 190

RESULT 67  
T02852  
probable membrane protein L1439.4 [imported] - Leishmania major (strain Friedlin)  
C:Species: Leishmania major  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 19-May-2000  
C:Accession: T02852; H81462  
R:Myler, P.J.  
submitted to the EMBL Data Library, May 1998  
A:Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.  
A:Reference number: Z14740  
A:Accession: T02852  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-800 <MYL>  
A:Cross-references: EMBL:AE001274; NID:g3264850; PID:g2266920  
R:Myler, P.J.; Audleman, L.; devos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;  
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999  
A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c  
A:Reference number: A81455; MUID:99178987; PMID:10077609  
A:Accession: H81462  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-800 <PYL>  
A:Cross-references: GB:AE001274; NID:g3264850; PIDN:AAC24675.1; PID:g2266920; GSPDB:GN0  
A:Experimental source: strain MHOM/IL/81/Friedlin  
C:Genetics:  
A:Gene: L1439.4  
A:Map position: 1

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 800;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDDG 444  
Db 21 DDDDDDDG 28

RESULT 68  
B96716  
probable serine/threonine kinase F23010.20 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: B96716  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
ansen, N.F.; Hughes, B.; Hutzar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: B96716  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-836 <STO>  
A:Cross-references: GB:AE005173; NID:g7705100; PIDN:AAF67779.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F23010.20  
A:Map position: 1

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 836;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 172 WMAPEVIO 179  
Db 412 WMAPEVIO 419

RESULT 69  
S48975  
hypothetical protein YHR131c - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 02-Dec-1994 #sequence\_revision 02-Dec-1994 #text\_change 19-Apr-2002  
C:Accession: S48975  
R:Fulton, L.  
submitted to the EMBL Data Library, June 1994  
A:Description: The sequence of S. cerevisiae cosmid 9315.  
A:Reference number: S48967  
A:Accession: S48975  
A:Molecule type: DNA  
A:Residues: 1-840 <FUL>  
A:Cross-references: EMBL:U10398; NID:g551328; PID:g500681; GSPDB:GN00008; MIPS:YHR131c  
C:Genetics:  
A:Gene: MIPS:YHR131c  
A:Cross-references: SGD:S0001173  
A:Map position: 8R

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 840;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 437 DDDDDDDG 444  
|||  
Db 829 DDDDDDDG 836

## RESULT 70

A53800  
mixed-lineage protein kinase (EC 2.7.1.-) 3 - human  
N/Alternate names: protein kinase PTK1; protein kinase SPRK  
C/Species: Homo sapiens (man)  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C/Accession: A53800; 158395  
R/Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.  
J. Biol. Chem. 269, 15092-15100, 1994  
A/Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont  
A/Reference number: A53800; MUID:94253068; PMID:8195146  
A/Accession: A53800  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-847 <GAL>  
A/Cross-references: GB:U07747; NID:G464027; PIDN:AAA19647.1; PID:G464028  
R/ing, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lassam, N.J.  
Oncogene 9, 1745-1750, 1994  
A/Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 doma  
A/Reference number: 158395; MUID:94239754; PMID:8183572  
A/Accession: 158395  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-847 <RES>  
A/Cross-references: GB:U32976; NID:G488295; PIDN:AAA59859.1; PID:G488296  
C/Genetics:  
A/Gene: GDB:MLK3; PTK1; SPRK  
A/Cross-references: GDB:134755; OMIM:600050  
A/Map position: 11q13.1-11q13.3  
C/Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology  
C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k  
F/48-100/Domain: SH3 homology <SH3>  
F/115-383/Domain: protein kinase homology <KIN>  
F/123-131/Region: protein kinase ATP-binding motif  
F/403-424/Region: leucine zipper motif  
F/438-459/Region: leucine zipper motif  
F/468-482/Region: basic

| Query Match             | 1.8%;   | Score 8;      | DB 1; | Length 847;       |
|-------------------------|---------|---------------|-------|-------------------|
| Best Local Similarity   | 100.0%; | Pred. No. 36; |       |                   |
| Matches 8; Conservative | 0;      | Mismatches    | 0;    | Indels 0; Gaps 0; |

Y 129 VIHRDLKS 136  
|||  
Db 237 VIHRDLKS 244

## RESULT 71

T18861  
probable peptidase-aspartate beta-dioxygenase (EC 1.14.11.16) - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C/Accession: T18861; T23533  
R/Swinburne, J.  
submitted to the EMBL Data Library, August 1996  
A/Reference number: Z19032  
A/Accession: T18861  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-872 <WIL>  
A/Cross-references: EMBL:Z79596; NID:el323798; PIDN:CAB01859.1; GSPDB:GN00028; CESP:K09A  
A/Experimental source: clone C02C6  
R/Swinburne, J.  
submitted to the EMBL Data Library, August 1996  
A/Reference number: Z19753  
A/Accession: T23533  
A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA  
A/Residues: 1-872 <WIL>  
A/Cross-references: EMBL:Z79601; PIDN:CAB01887.1; GSPDB:GN00028; CESP:K09A9.6  
A/Experimental source: clone K09A9  
C/Genetics:  
A/Gene: CESP:K09A9.6  
A/Map position: X  
A/Introns: 18/1; 40/3; 74/1; 104/2; 139/1; 170/3; 385/3; 432/2; 464/3; 530/2; 609/3; 776  
C/Keywords: oxidoreductase

| Query Match             | 1.8%;   | Score 8;      | DB 2; | Length 872;       |
|-------------------------|---------|---------------|-------|-------------------|
| Best Local Similarity   | 100.0%; | Pred. No. 37; |       |                   |
| Matches 8; Conservative | 0;      | Mismatches    | 0;    | Indels 0; Gaps 0; |

QY 436 GDDDDDDDD 443  
|||  
Db 288 GDDDDDDDD 295

## RESULT 72

A56731  
chromatin assembly factor I p150 chain - human  
C/Species: Homo sapiens (man)  
C/Date: 08-Feb-1996 #sequence\_revision 08-Feb-1996 #text\_change 21-Jul-2000  
C/Accession: A56731  
R/Kaufman, P.D.; Kobayashi, R.; Kessler, N.; Stillman, B.  
Cell 81, 1105-1114, 1995  
A/Title: The p150 and p60 subunits of chromatin assembly factor I: a molecular link bet  
A/Reference number: A56731; MUID:95323966; PMID:7600578  
A/Accession: A56731  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-938 <KAU>  
A/Cross-references: GB:U20979; NID:G882257; PIDN:AAA76736.1; PID:G882258

| Query Match             | 1.8%;   | Score 8;      | DB 2; | Length 938;       |
|-------------------------|---------|---------------|-------|-------------------|
| Best Local Similarity   | 100.0%; | Pred. No. 40; |       |                   |
| Matches 8; Conservative | 0;      | Mismatches    | 0;    | Indels 0; Gaps 0; |

QY 434 SEGDDDDDD 441  
|||  
Db 598 SEGDDDDDD 605

## RESULT 73

B45082  
neurotrophic receptor ror2 precursor - human  
N/Contains: protein-tyrosine kinase (EC 2.7.1.112)  
C/Species: Homo sapiens (man)  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 19-May-2000  
C/Accession: B45082  
R/Masiakowski, P.; Carroll, R.D.  
J. Biol. Chem. 267, 26181-26190, 1992  
A/Title: A novel family of cell surface receptors with tyrosine kinase-like domain.  
A/Reference number: A45082; MUID:93100347; PMID:1334494  
A/Accession: B45082  
A/Molecule type: mRNA  
A/Residues: 1-943 <MAS>  
A/Cross-references: GB:M97639; NID:G337466; PIDN:AAA60276.1; PID:G337467  
A/Note: sequence extracted from NCBI backbone (NCBIP:120918)  
C/Genetics:  
A/Gene: GDB:NTRKR2  
A/Cross-references: GDB:136454  
A/Map position: 6p21-6p21  
C/Superfamily: neurotrophic receptor ror; immunoglobulin homology; kringle homology; pr  
C/Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyro  
F/1-27/Domain: signal sequence #status predicted <SIG>  
F/28-937/Product: neurotrophic receptor ror2 #status predicted <MAT>  
F/76-137/Domain: immunoglobulin homology <IMM>  
F/316-394/Domain: kringle homology <KR>  
F/412-428/Domain: transmembrane #status predicted <TMN>  
F/471-753/Domain: protein kinase homology <KIN>  
F/479-487/Region: protein kinase ATP-binding motif

F:70,188,318/Binding site: carbohydrate (Asn) (covalent) #status predicted

## Query Match

Best Local Similarity 1.8%; Score 8; DB 2; Length 943;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVLE 198  
DB 677 SYGVVLE 684

## RESULT 74

T05335

hypothetical protein F1C12.190 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 20-Sep-1999

C/Accession: T05335  
R/Bevan, M.; Terry, N.; Ardiles, W.; Buyssehaert, C.; Dasseville, R.; De Clerck, R.; De

es, H.W.; Mayer, K.F.X.; Schueller, C.  
Submitted to the Protein Sequence Database, April 1998

A/Reference number: 215408  
A/Accession: T05335

A/Molecule type: DNA

A/Residues: 1-992 <BEV>

A/Cross-references: EMBL:AL022224

A/Experimental source: cultivar Columbia; BAC clone F1C12

C/Genetics:

A/Map position: 4

A/Introns: 884/1

A/Note: F1C12.190

C/Superfamily: protein kinase Xa21, leucine-rich alpha-2-glycoprotein repeat homology; F

## Query Match

Best Local Similarity 1.8%; Score 8; DB 2; Length 992;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 EVAVKKLL 48  
DB 734 EVAVKKLL 741

## RESULT 75

S49835

hypothetical protein YDR080w - yeast (Saccharomyces cerevisiae)  
N/Alternate names: hypothetical protein D4446; hypothetical protein YD8554.13

C/Species: Saccharomyces cerevisiae

C/Date: 13-Jan-1995 #sequence\_revision 10-Feb-1995 #text\_change 29-Oct-1999  
A/Accession: S49835; S48767; S55827; S67897

R/Richards, C.; Harris, D.E.  
Submitted to the EMBL Data Library, November 1994

A/Reference number: S49835  
A/Accession: S49835

A/Molecule type: DNA

A/Residues: 1-992 <RIC>

A/Cross-references: EMBL:Z46796; NID:G577794; PIDN:CAA86802.1; PID:G577807  
R/Coster, F.; Jonniaux, J.L.; Goffeau, A.  
Submitted to the EMBL Data Library, October 1994

A/Reference number: S48758  
A/Accession: S48758

A/Molecule type: DNA

A/Residues: 1-423, 'M', 425-992 <COS>

A/Cross-references: EMBL:X82086; NID:G558241; PIDN:CAA57607.1; PID:G558251  
R/Coster, F.; Jonniaux, J.L.; Goffeau, A.  
Yeast 11, 673-679, 1995

A/Title: Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open reading fr

A/Reference number: S55819; MUID:96093910; PMID:7483840  
A/Accession: S55827

A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA

A/Residues: 1-423, 'M', 425-992 <COM>

A/Cross-references: EMBL:X82086; NID:G558241; PIDN:CAA57607.1; PID:G558251

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994  
R/Foury, F.; Jonniaux, J.L.; Purnelle, B.; Coster, F.; Goffeau, A.

Submitted to the Protein Sequence Database, July 1996  
A/Reference number: S67889

A/Accession: S67897

A/Molecule type: DNA

A/Residues: 1-992 <FOU>

A/Cross-references: EMBL:Z74376; NID:G1431548; PIDN:CAA98899.1; PID:e253331; PID:G14315  
A/Experimental source: strain S288C

C/Genetics:

A/Gene: SGD:VPS41

A/Cross-references: SGD:S0002487; MIPS:YDR080w

A/Map position: 4R

## Query Match

Best Local Similarity 1.8%; Score 8; DB 2; Length 992;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
DB 78 GDDDDDD 85

Search completed: May 1, 2003, 20:56:24  
Job time: 73 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:39:47 ; Search time 26 Seconds  
(without alignments)  
725.836 Million cell updates/sec

Title: US-09-757-982-5  
Perfect score: 455  
Sequence: 1 MSSSLGASRVQIKFDLDQFFE.....GDDDDDDGEEEDNDMDNSE 455

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues  
d size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID           | Description         |
|------------|-------|-------------|--------|--------------|---------------------|
| 1          | 10    | 2.2         | 729    | 1 KEX1_YEAST | P09620 saccharomyc  |
| 2          | 10    | 2.2         | 1089   | 1 NMD2_YEAST | P38798 saccharomyc  |
| 3          | 9     | 2.0         | 375    | 1 MCK1_YEAST | P21965 saccharomyc  |
| 4          | 9     | 2.0         | 502    | 1 KIR3_MOUSE | Q61288 mus musculu  |
| 5          | 9     | 2.0         | 505    | 1 KIR3_MOUSE | P80203 rattus norv  |
| 6          | 9     | 2.0         | 1363   | 1 ILPR_BRALA | O02466 brachioisto  |
| 7          | 9     | 2.0         | 1390   | 1 INSR_AEDAE | O93105 aedes aegypt |
| 8          | 8     | 1.8         | 157    | 1 UL51_HCMVA | P16792 human cytom  |
| 9          | 8     | 1.8         | 158    | 1 2SS_SOYBN  | P19594 glycine max  |
| 10         | 8     | 1.8         | 229    | 1 REGQ_BP82  | P13870 bacterioph   |
| 11         | 8     | 1.8         | 286    | 1 MCM1_YEAST | P11746 saccharomyc  |
| 12         | 8     | 1.8         | 294    | 1 OSTP_MOUSE | P10923 mus musculu  |
| 13         | 8     | 1.8         | 317    | 1 OSTP_MOUSE | P08721 rattus norv  |
| 14         | 8     | 1.8         | 352    | 1 RLA0_HALCU | P17006 halobacteri  |
| 15         | 8     | 1.8         | 352    | 1 RLA0_HALCU | P13553 halobacteri  |
| 16         | 8     | 1.8         | 391    | 1 Y532_METJA | Q57952 methanococc  |
| 17         | 8     | 1.8         | 482    | 1 T2EA_YEAST | P36100 saccharomyc  |
| 18         | 8     | 1.8         | 487    | 1 STR4_HUMAN | Q13043 homo sapien  |
| 19         | 8     | 1.8         | 491    | 1 STR3_HUMAN | Q13188 homo sapien  |
| 20         | 8     | 1.8         | 579    | 1 M3K7_MOUSE | Q62073 mus musculu  |
| 21         | 8     | 1.8         | 590    | 1 YM72_YEAST | Q05021 saccharomyc  |
| 22         | 8     | 1.8         | 606    | 1 M3K7_HUMAN | Q43318 homo sapien  |
| 23         | 8     | 1.8         | 746    | 1 EZH2_MOUSE | Q15910 homo sapien  |
| 24         | 8     | 1.8         | 746    | 1 EZH2_MOUSE | Q61188 mus musculu  |
| 25         | 8     | 1.8         | 799    | 1 CN3B_MOUSE | Q61409 mus musculu  |
| 26         | 8     | 1.8         | 840    | 1 YHT1_YEAST | P38835 saccharomyc  |
| 27         | 8     | 1.8         | 875    | 1 DD10_HUMAN | Q13206 homo sapien  |
| 28         | 8     | 1.8         | 938    | 1 CAFA_HUMAN | Q13111 homo sapien  |
| 29         | 8     | 1.8         | 943    | 1 ROR2_HUMAN | Q01974 homo sapien  |
| 30         | 8     | 1.8         | 944    | 1 ROR2_MOUSE | Q92138 mus musculu  |
| 31         | 8     | 1.8         | 992    | 1 VP41_YEAST | P38959 saccharomyc  |
| 32         | 8     | 1.8         | 1085   | 1 IFH1_YEAST | P39520 saccharomyc  |
| 33         | 8     | 1.8         | 1131   | 1 YAB9_YEAST | P31380 saccharomyc  |

|     |   |     |      |              |                    |
|-----|---|-----|------|--------------|--------------------|
| 34  | 8 | 1.8 | 1165 | 1 YNFA_YEAST | P53950 saccharomyc |
| 35  | 7 | 1.5 | 86   | 1 RNPH_AZOVI | Q9f5y0 azotobacter |
| 36  | 7 | 1.5 | 101  | 1 KOA2_ECOLI | P03052 escherichia |
| 37  | 7 | 1.5 | 147  | 1 UCRH_YEAST | P00127 saccharomyc |
| 38  | 7 | 1.5 | 160  | 1 HPEK_HAEIN | P43777 haemophilus |
| 39  | 7 | 1.5 | 172  | 1 UBC2_YEAST | P06104 saccharomyc |
| 40  | 7 | 1.5 | 173  | 1 NUGM_BRABR | P43194 brachyramph |
| 41  | 7 | 1.5 | 173  | 1 NUGM_BRAMA | P43195 brachyramph |
| 42  | 7 | 1.5 | 175  | 1 VIRR_AGRTR | O52278 agrobacteri |
| 43  | 7 | 1.5 | 183  | 1 NO29_XENLA | O42584 xenopus lae |
| 44  | 7 | 1.5 | 204  | 1 HMG1_ONCMY | P07746 oncorhynch  |
| 45  | 7 | 1.5 | 207  | 1 H1T_MOUSE  | Q07133 mus musculu |
| 46  | 7 | 1.5 | 228  | 1 EF1B_ORYSA | Q40680 oryza sativ |
| 47  | 7 | 1.5 | 230  | 1 EF1B_BETVU | O81918 beta vulgar |
| 48  | 7 | 1.5 | 245  | 1 ICP3_HSV1N | P37319 herpes sim  |
| 49  | 7 | 1.5 | 248  | 1 ICP3_HSV11 | P36313 herpes sim  |
| 50  | 7 | 1.5 | 250  | 1 YB95_ARATH | O8w487 arabidopsi  |
| 51  | 7 | 1.5 | 252  | 1 ICP3_HSV1D | P37318 herpes sim  |
| 52  | 7 | 1.5 | 253  | 1 ASP_PLAFL  | P13825 plasmodium  |
| 53  | 7 | 1.5 | 253  | 1 L139_CAEEL | P34684 caenorhabdi |
| 54  | 7 | 1.5 | 253  | 1 T2D7_RAT   | Q62880 rattus norv |
| 55  | 7 | 1.5 | 263  | 1 ICP3_HSV1F | P08353 herpes sim  |
| 56  | 7 | 1.5 | 264  | 1 OSTP_CHICK | P23498 gallus gall |
| 57  | 7 | 1.5 | 264  | 1 T2D7_HUMAN | Q16594 homo sapien |
| 58  | 7 | 1.5 | 265  | 1 EF1D_XENLA | P29693 xenopus lae |
| 59  | 7 | 1.5 | 280  | 1 HAX1_MOUSE | O35387 mus musculu |
| 60  | 7 | 1.5 | 289  | 1 YNV9_CAEEL | P34572 caenorhabdi |
| 61  | 7 | 1.5 | 295  | 1 UBC3_YEAST | P34572 caenorhabdi |
| 62  | 7 | 1.5 | 296  | 1 PEX5_CANAL | O74711 candida alb |
| 63  | 7 | 1.5 | 303  | 1 MURB_BACSU | P18579 bacillus su |
| 64  | 7 | 1.5 | 305  | 1 CDK3_HUMAN | O00526 homo sapien |
| 65  | 7 | 1.5 | 309  | 1 MLF_DROME  | O9nkvo drosophila  |
| 66  | 7 | 1.5 | 310  | 1 DOS2_YEAST | P54858 saccharomyc |
| 67  | 7 | 1.5 | 314  | 1 MK03_MOUSE | Q63844 m mitogen-a |
| 68  | 7 | 1.5 | 323  | 1 KRAF_MS36  | P00532 murine sarc |
| 69  | 7 | 1.5 | 323  | 1 LUKE_STAAN | P31715 staphylococ |
| 70  | 7 | 1.5 | 325  | 1 H1GB_STAAN | Q07226 staphylococ |
| 71  | 7 | 1.5 | 328  | 1 KRAB_MOUSE | P28028 mus musculu |
| 72  | 7 | 1.5 | 332  | 1 AAIP_WHEAT | Q02066 triticum ae |
| 73  | 7 | 1.5 | 332  | 1 CDC2_CAEEL | P34556 caenorhabdi |
| 74  | 7 | 1.5 | 332  | 1 Y713_CHLMU | Q9pjw3 chlamydia m |
| 75  | 7 | 1.5 | 333  | 1 SPT2_YEAST | P06843 saccharomyc |
| 76  | 7 | 1.5 | 348  | 1 RLA0_HALMA | P15825 halocarcula |
| 77  | 7 | 1.5 | 349  | 1 STY1_SCHPO | O09892 schizosacch |
| 78  | 7 | 1.5 | 353  | 1 ASK2_ARATH | P43292 arabidopsi  |
| 79  | 7 | 1.5 | 353  | 1 FUS3_YEAST | P16892 saccharomyc |
| 80  | 7 | 1.5 | 355  | 1 ARG3_YEAST | P07250 saccharomyc |
| 81  | 7 | 1.5 | 358  | 1 MK01_MOUSE | P27703 mus musculu |
| 82  | 7 | 1.5 | 359  | 1 PKX1_HUMAN | P18187 homo sapien |
| 83  | 7 | 1.5 | 359  | 1 KMOS_XENLA | P12965 xenopus lae |
| 84  | 7 | 1.5 | 360  | 1 MK01_BOVIN | P46196 bos taurus  |
| 85  | 7 | 1.5 | 360  | 1 MK01_HUMAN | P28482 homo sapien |
| 86  | 7 | 1.5 | 361  | 1 MK01_XENLA | P26696 xenopus lae |
| 87  | 7 | 1.5 | 363  | 1 ASK1_ARATH | P43291 arabidopsi  |
| 88  | 7 | 1.5 | 365  | 1 M14B_DROME | O61443 drosophila  |
| 89  | 7 | 1.5 | 366  | 1 M14A_DROME | O62618 drosophila  |
| 90  | 7 | 1.5 | 367  | 1 MK12_MOUSE | O08911 mus musculu |
| 91  | 7 | 1.5 | 367  | 1 MK12_MOUSE | Q63538 rattus norv |
| 92  | 7 | 1.5 | 367  | 1 RMIL_AV111 | Q10533 avian retro |
| 93  | 7 | 1.5 | 368  | 1 MPK7_ARATH | Q39027 arabidopsi  |
| 94  | 7 | 1.5 | 370  | 1 CTP1_PLAFL | P49587 plasmodium  |
| 95  | 7 | 1.5 | 370  | 1 MPK1_ARATH | Q39021 arabidopsi  |
| 96  | 7 | 1.5 | 370  | 1 MPK3_ARATH | Q39023 arabidopsi  |
| 97  | 7 | 1.5 | 370  | 1 YG2G_YEAST | P53249 saccharomyc |
| 98  | 7 | 1.5 | 371  | 1 NTF6_TOBAC | Q40531 nicotiana t |
| 99  | 7 | 1.5 | 372  | 1 NTF3_TOBAC | Q40517 nicotiana t |
| 100 | 7 | 1.5 | 373  | 1 YFJ4_YEAST | P43603 saccharomyc |

ALIGNMENTS

RESULT 1



KEX1\_YEAST STANDARD; PRT; 729 AA.  
 ID KEX1\_YEAST  
 AC P09620;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Carboxypeptidase KEX1 precursor (EC 3.4.16.6) (Carboxypeptidase D).  
 GN KEX1 OR YGL203C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NCBI\_TaxID=4932;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87273520; PubMed=3301004;  
 RA Dmochowska A., Dignard D., Henning D., Thomas D.Y., Bussey H.;  
 RT "Yeast KEX1 gene encodes a putative protease with a carboxypeptidase  
 B-like function involved in killer toxin and alpha-factor precursor  
 processing.";  
 RT Cell 50:573-584(1987).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX Bjourson A.J., McReynolds A.D.K., Wright L.F.;  
 RA Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.  
 [3]  
 RP CHARACTERIZATION.  
 RX MEDLINE=89343988; PubMed=2668738;  
 RA Cooper A., Bussey H.;  
 RT "Characterization of the yeast KEX1 gene product: a carboxypeptidase  
 involved in processing secreted precursor proteins.";  
 RT Mol. Cell. Biol. 9:2706-2714(1989).  
 [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=96276234; PubMed=8745419;  
 RA Shilton B.H., Li Y., Tessier D., Thomas D.Y., Cygler M.;  
 RT "Crystallization of a soluble form of the Kexlp serine  
 carboxypeptidase from Saccharomyces cerevisiae.";  
 RT Protein Sci. 5:395-397(1996).  
 CC -!- FUNCTION: PROTEASE WITH A CARBOXYPEPTIDASE B-LIKE FUNCTION  
 INVOLVED IN KILLER TOXIN AND ALPHA-FACTOR PRECURSOR PROCESSING.  
 CC -!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine  
 or lysine residue.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M17231; AAA34717.1; -.  
 DR EMBL; Z72725; CAA96915.1; -.  
 DR PIR; A29651; A29651.  
 DR PDB; 1AC5; 15-MAY-97.  
 DR MEROPS; S10.007; -.  
 DR SGD; S0003171; KEX1.  
 DR InterPro; IPR000379; Ser esters site.  
 DR InterPro; IPR001563; Serine catpept.  
 DR Pfam; PF00450; serine\_catpept; 1.  
 DR PRINTS; PR00724; CRBOXYPTASEC.  
 DR ProDom; PD001189; Serine catpept; 1.  
 DR PROSITE; PS00131; CARBOXYPEPT\_SER\_SER; 1.  
 DR PROSITE; PS00560; CARBOXYPEPT\_SER\_HIS; 1.  
 KM Hydrolase; Carboxypeptidase; Glycoprotein; Transmembrane; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 22  
 FT CHAIN 23 729 CARBOXYPEPTIDASE KEX1.  
 FT ACT\_SITE 198 198  
 FT ACT\_SITE 405 405 BY SIMILARITY.  
 FT ACT\_SITE 470 470 BY SIMILARITY.  
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 467 467 N-LINKED (GLCNAC. . .).  
 FT DOMAIN 506 611 ASP/GLU-RICH (ACIDIC).  
 FT TRANSMEM 619 637 POTENTIAL.  
 FT MUTAGEN 198 198 S->A: INACTIVATES ENZYME.  
 SQ SEQUENCE 729 AA; 82245 MW; 70583F279AC02A41 CRC64;  
 Query Match 2.2%; Score 10; DB 1; Length 729;  
 Best Local Similarity 100.0%; Pred. No. 0.092;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 434 SEGDDDDDDDD 443  
 Db 564 SEGDDDDDDDD 573  
 RESULT 2  
 NMD2\_YEAST STANDARD; PRT; 1089 AA.  
 ID NMD2\_YEAST  
 AC P38798;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Nonsense-mediated mRNA decay protein 2 (Up-frameshift suppressor 2).  
 GN NMD2 OR UPF2 OR IFS1 OR SUV1 OR YHR077C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NCBI\_TaxID=4932;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95189083; PubMed=7883168;  
 RA He F., Jacobson A.;  
 RT "Identification of a novel component of the nonsense-mediated mRNA  
 decay pathway by use of an interacting protein screen.";  
 RT Genes Dev. 9:437-454(1995).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PLY136;  
 RX MEDLINE=95189082; PubMed=7883167;  
 RA Cui Y., Hagan K.W., Zhang S., Peltz S.W.;  
 RT "Identification and characterization of genes that are required for  
 the accelerated degradation of mRNAs containing a premature  
 translational termination codon.";  
 RT Genes Dev. 9:423-436(1995).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95327692; PubMed=7604038;  
 RA Lee S.I., Umen J.G., Varmus H.E.;  
 RT "A genetic screen identifies cellular factors involved in retroviral  
 -1 frameshifting.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 92:6587-6591(1995).  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=94378003; PubMed=8091229;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,  
 Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,  
 Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,  
 Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,  
 Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,  
 Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,  
 Vaudin M.;  
 RA "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome  
 VII.";  
 RT Science 265:2077-2082(1994).  
 CC -!- FUNCTION: INVOLVED IN NONSENSE-MEDIATED DECAY OF MRNAS CONTAINING  
 CC PREMATURE STOP CODONS. IT INTERACTS, VIA ITS C-TERMINUS, WITH  
 CC NAM7/UPF1. COULD BE INVOLVED IN DETERMINING THE EFFICIENCY OF  
 CC TRANSLATIONAL TERMINATION OR REINITIATION OR FACTORS INVOLVED IN  
 CC THE INITIAL ASSEMBLY OF AN INITIATION- AND TERMINATION-COMPETENT  
 CC MRNP.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; U14974; AAA67724.1; -  
DR EMBL; U12137; AAA66521.1; -  
DR EMBL; U28158; AAA74948.1; -  
DR EMBL; U10556; AAB6893.1; -  
DR PIR; S46815; S46815.  
DR SGD; S0001119; NMD2.  
DR InterPro; IPR003890; IF\_eIF4G.  
DR Pfam; PF02854; MIF4G; 3.  
DR SMART; SM00543; MIF4G; 3.  
FT DOMAIN 843 975 ASP/GLU-RICH (HIGHLY ACIDIC).  
FT CONFLICT 2 2 D -> YQQ (IN REF. 3 AND 4).  
SEQUENCE 1089 AA; 126746 MW; 13BBE725675C8F52 CRC64;

Query Match 2.2%; Score 10; DB 1; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 437 DDDDDDDGEE 446  
Db 902 DDDDDDDGEE 911

## RESULT 3

MCK1\_YEAST  
ID MCK1\_YEAST STANDARD; PRT; 375 AA.  
AC P21965;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Protein kinase MCK1 (EC 2.7.1.-) (Meiosis and centromere regulatory  
kinase).  
GN MCK1 OR YPK1 OR YNL307C OR N0392.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces;  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
RX MEDLINE=91061730; PubMed=1701015;  
RT Dailey D., Schieven G.L., Lim M.-Y., Margardt H., Gilmore T.,  
RT Thorne J., Martin G.S.;  
RT "Novel yeast protein kinase (YPK1 gene product) is a 40-kilodalton  
RT phosphotyrosyl protein associated with protein-tyrosine kinase  
RT activity.";  
RL Mol. Cell. Biol. 10:6244-6256(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91184610; PubMed=2010083;  
RT Neigelborn L., Mitchell A.P.;  
RT "The yeast MCK1 gene encodes a protein kinase homolog that activates  
RT early meiotic gene expression.";  
RL Genes Dev. 5:533-548(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91184611; PubMed=2010084;  
RT Shero J.H., Hietter P.;  
RT "A suppressor of a centromere DNA mutation encodes a putative protein  
RT kinase (MCK1).";  
RL Genes Dev. 5:549-560(1991).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX STRAIN=S288C / FY1676;  
RX MEDLINE=96076632; PubMed=7502583;  
RA Malfah M., Nicaud J.-M., Levesque H., Gaillardin C.;

RT "Sequencing analysis of a 24.7 kb fragment of yeast chromosome XIV  
RT identifies six known genes, a new member of the hexose transporter  
RT family and ten new open reading frames.";  
RL Yeast 11:1077-1085(1995).  
RN [5]

RP SEQUENCE OF 1-16 FROM N.A.  
RX STRAIN=S288C / FY1679;  
RX MEDLINE=96132033; PubMed=8553702;  
RT Maurer K.C.T., Urbanus J.H.M., Planta R.J.;  
RT "Sequence analysis of a 30 kb DNA segment from yeast chromosome XIV  
RT carrying a ribosomal protein gene cluster, the genes encoding a  
RT plasma membrane protein and a subunit of replication factor C, and a  
RT novel putative serine/threonine protein kinase gene.";  
RL Yeast 11:1303-1310(1995).  
CC -1- FUNCTION: MAY BE AN AUTOPHOSPHORYLATING TYROSINE KINASE, A  
CC BIFUNCTIONAL (SERINE/TYROSINE-SPECIFIC) PROTEIN KINASE, OR A  
CC SERINE KINASE THAT IS A SUBSTRATE FOR AN ASSOCIATED TYROSINE  
CC KINASE. MCK1 IS A TRANSCRIPTIONAL ACTIVATOR OF IME1, IT STIMULATES  
CC SPORE MATURATION, AND PLAY A POSITIVE REGULATORY ROLE IN BOTH  
CC MITOTIC CENTROMERE FUNCTION AND ACTIVATION OF EARLY MEIOTIC GENE  
CC EXPRESSION.  
CC -1- PTM: PHOSPHORYLATED AT TYROSINE AND SERINE IN VIVO.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; X55054; CAA38895.1; -  
DR EMBL; M55984; AAA34764.1; -  
DR EMBL; Z46259; CAA86388.1; -  
DR EMBL; Z71583; CAA96236.1; -  
DR EMBL; Z71582; CAA96235.1; -  
DR PIR; A36354; A36354.  
DR PIR; A39622; A39622.  
DR PIR; A39623; A39623.  
DR HSSP; P24941; 1A01.  
DR SGD; S0005251; MCK1.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding;  
KW phosphorylation.  
FT SIMILAR 1 8 TO CAMP-DEPENDENT PROTEIN KINASE TYPE 1  
FT (AA 1-8) (IDENTICAL).  
FT DOMAIN 35 327 PROTEIN KINASE.  
FT NP\_BIND 41 49 ATP (BY SIMILARITY).  
FT BINDING 68 68 ATP (BY SIMILARITY).  
FT ACT\_SITE 164 164 BY SIMILARITY.  
FT MOD\_RES 199 199 PHOSPHORYLATION (BY SIMILARITY).  
SQ SEQUENCE 375 AA; 43136 MW; BCF0C776B6E3841B CRC64;

Query Match 2.0%; Score 9; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 0.51;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 145 GVLKICDFG 153  
Db 177 GVLKICDFG 185

RESULT 4  
KIR3\_MOUSE  
ID KIR3\_MOUSE STANDARD; PRT; 502 AA.

AC Q61288; Q61289; 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Serine/threonine-protein kinase receptor R3 precursor (EC 2.7.1.37)  
 DE (SKR3) (Activin receptor-like kinase 1) (ALK-1) (TGF-B superfamily  
 DE receptor type I) (TSR-I).  
 GN ACVRL1 OR ACVRLK1 OR ALK-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=96067535; PubMed=7488127;  
 RA Wu X., Robinson C.E., Fong H.W., Crabtree J.S., Rodriguez B.R.,  
 RA Roe B.A., Gimble J.M.;  
 RA "Cloning and characterization of the murine activin receptor like  
 RT kinase-1 (ALK-1) homolog."  
 RT Biochem. Biophys. Res. Commun. 216:78-83(1995).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=95269711; PubMed=7750489;  
 RA Dewulf N., Verschueren K., Lonnoy O., Moren A., Grimsby S.,  
 RA Spiegler K., Miyazono K., Huybrechts D., ten Dijke P.;  
 RT "Distinct spatial and temporal expression patterns of two type I  
 RT receptors for bone morphogenetic proteins during mouse  
 RT embryogenesis."  
 RT Endocrinology 136:2652-2663(1995).  
 CC -1- FUNCTION: TYPE I/TYPE II TGF-BETA RECEPTORS FORM AN HETEROMERIC  
 CC COMPLEX AFTER BINDING TGF-BETA AT THE CELL SURFACE AND ACT AS  
 CC SIGNAL TRANSDUCERS. MAY BIND ACTIVIN AS WELL.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC TGF-B RECEPTOR SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L48015; AAB03642.1; -;  
 CC EMBL; Z31664; CAA83484.1; -;  
 CC MGI; MGI:1338946; Acvrl1.  
 DR InterPro; IPR000472; Activin\_rec.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR003605; TGFbeta\_GS.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF01064; Activin\_rec; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00467; GS; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR Receptor; Transferase; Serine/threonine-protein kinase; ATP-binding;  
 KW Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 502  
 FT R3.  
 FT DOMAIN 23 119  
 FT TRANSMEM 120 140  
 FT DOMAIN 141 502  
 FT DOMAIN 201 502  
 FT NP\_BIND 208 215  
 FT BINDING 228 228  
 FT ACT\_SITE 329 329  
 FT BY SIMILARITY.

FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 17 17 F -> L (IN REF. 2).  
 FT CONFLICT 21 21 R -> Q (IN REF. 2).  
 FT CONFLICT 23 23 D -> R (IN REF. 2).  
 FT CONFLICT 25 26 RR -> AK (IN REF. 2).  
 FT CONFLICT 305 305 A -> P (IN REF. 2).  
 FT CONFLICT 358 359 SD -> NE (IN REF. 2).  
 FT CONFLICT 366 366 N -> T (IN REF. 2).  
 SQ SEQUENCE 502 AA; 56632 MW; 092ABB29778BB004 CRC64;  
 Query Match 2.0%; Score 9; DB 1; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 0.67;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 131 HRDLKSRNV 139  
 Db 327 HRDLKSRNV 335  
 ID KIR3\_RAT STANDARD; PRT; 505 AA.  
 KIR3\_RAT  
 ID KIR3\_RAT  
 AC P80203;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Serine/threonine-protein kinase receptor R3 precursor (EC 2.7.1.37)  
 DE (SKR3) (TGF-B superfamily receptor type I) (TSR-I).  
 GN ACVRL1 OR ACVRLK1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Urogenital ridge;  
 RX MEDLINE=93372378; PubMed=8395914;  
 RA He W.-W., Gustafson M., Hirobe S., Donahoe P.;  
 RA "Developmental expression of four novel serine/threonine kinase  
 RT receptors homologous to the activin/transforming growth factor-beta  
 RT type II receptor family."  
 RT Dev. Dyn. 196:133-142(1993).  
 CC -1- FUNCTION: TYPE I/TYPE II TGF-BETA RECEPTORS FORM AN HETEROMERIC  
 CC COMPLEX AFTER BINDING TGF-BETA AT THE CELL SURFACE AND ACT AS  
 CC SIGNAL TRANSDUCERS. MAY BIND ACTIVIN AS WELL.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: UROGENITAL RIDGE, TESTIS, OVARY, BRAIN AND  
 CC LUNGS.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC TGF-B RECEPTOR SUBFAMILY.  
 CC InterPro; IPR000472; Activin\_rec.  
 CC InterPro; IPR000719; Euk\_pkinase.  
 CC InterPro; IPR002290; Ser\_thr\_pkinase.  
 CC InterPro; IPR003605; TGFbeta\_GS.  
 CC Pfam; PF00069; pkinase; 1.  
 CC Pfam; PF01064; Activin\_rec; 1.  
 CC ProDom; PD000001; Euk\_pkinase; 1.  
 CC SMART; SM00467; GS; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC Receptor; Transferase; Serine/threonine-protein kinase; ATP-binding;  
 KW Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 20  
 FT CHAIN 21 505  
 FT R3.  
 FT DOMAIN 21 121  
 FT TRANSMEM 122 142  
 FT DOMAIN 143 505  
 FT DOMAIN 203 505  
 FT NP\_BIND 209 217  
 FT ATP (BY SIMILARITY).



FT BINDING 230 230 ATP (BY SIMILARITY).  
FT ACT\_SITE 332 332 BY SIMILARITY.  
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 505 AA; 56774 MW; 4C10D09C9105F23C CRC64;  
Query Match 2.0%; Score 9; DB 1; Length 505;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
Db 330 HRDLKSRNV 338

RESULT 6  
ILPR\_BRLA  
ID ILPR\_BRLA STANDARD; PRT; 1363 AA.  
AC 002466;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Insulin-like peptide receptor precursor (EC 2.7.1.112) (ILP receptor).  
OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus).  
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
OC Branchiostoma.  
OX NCBI\_TaxID=7740;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96408719; PubMed=8813726;  
RA Pashmforoush M., Chan S.J., Steiner D.F.;  
RT "Structure and expression of the insulin-like peptide receptor from amphioxus."  
RL Mol. Endocrinol. 10:857-866 (1996).  
CC -1- FUNCTION: THIS RECEPTOR BINDS TO THE INSULIN RELATED PEPTIDE AND HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.  
CC -1- SUBUNIT: PROBABLE Tetramer of 2 alpha and 2 beta chains linked by disulfide bonds. The alpha chains contribute to the formation of the ligand-binding domain, while the beta chain carry the kinase domain (by similarity).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).  
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN RECEPTOR SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; S83394; AAB50848.1; .  
CC HSP; P06213; IIRK.  
DR InterPro; IPR000494; EGFR\_L domain.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR003962; FNIII repeat.  
DR InterPro; IPR002174; Furin-like.  
DR InterPro; IPR002011; RTkinaseII.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00041; fn3; 3.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00757; Furin-like; 1.  
DR Pfam; PF01030; Recep\_L domain; 2.  
DR PRINTS; PR00014; FNTYPEIIIT.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00060; FN3; 3.  
DR SMART; SM00261; FU; 1.  
DR SMART; SM00219; TyrKc; 1.

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
KW Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;  
KW Glycoprotein; ATP-binding; Phosphorylation; Signal.  
FT SIGNAL 1 29  
FT CHAIN 30 716  
FT PROPEP 717 720  
FT CHAIN 721 1363  
FT DOMAIN 721 928  
FT TRANSMEM 929 949  
FT DOMAIN 950 1363  
FT DOMAIN 994 1283  
FT NP\_BIND 1000 1008  
FT BINDING 1028 1028  
FT ACT\_SITE 1148 1148  
FT MOD\_RES 1174 1174  
FT CARBOHYD 51 51  
FT CARBOHYD 97 97  
FT CARBOHYD 137 137  
FT CARBOHYD 278 278  
FT CARBOHYD 483 483  
FT CARBOHYD 599 599  
FT CARBOHYD 617 617  
FT CARBOHYD 665 665  
FT CARBOHYD 666 666  
FT CARBOHYD 711 711  
FT CARBOHYD 732 732  
FT CARBOHYD 736 736  
FT CARBOHYD 743 743  
FT CARBOHYD 816 816  
FT CARBOHYD 885 885  
FT CARBOHYD 898 898  
SQ SEQUENCE 1363 AA; 154104 MW; 238120B4EAB1ED65 CRC64;

Query Match 2.0%; Score 9; DB 1; Length 1363;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVWLWEM 199  
Db 1210 SYGVWLWEM 1218

RESULT 7  
INSR\_AEDAE  
ID INSR\_AEDAE STANDARD; PRT; 1390 AA.  
AC Q93105;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Insulin-like receptor precursor (EC 2.7.1.112) (MIR).  
OS Aedes aegypti (Yellowfever mosquito).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;  
OC Culicoidae; Aedes.  
OX NCBI\_TaxID=7159;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=UGAL; TISSUE=Ovary;  
RX MEDLINE=97254344; PubMed=9099579;  
RA Graf R., Neuenschwander S., Brown M.R., Ackermann U.;  
RT "Insulin-mediated secretion of ecdysteroids from mosquito ovaries and molecular cloning of the insulin receptor homologue from ovaries of blooded Aedes aegypti."  
RT Insect Mol. Biol. 6:151-163 (1997).  
CC -1- FUNCTION: THIS RECEPTOR PROBABLY BINDS AN INSULIN RELATED PROTEIN AND HAS A TYROSINE-PROTEIN KINASE ACTIVITY (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.



CC -1- SUBUNIT: TETRAMER OF 2 ALPHA AND 2 BETA CHAINS LINKED BY DISULFIDE  
CC BONDS. THE ALPHA CHAINS CONTRIBUTE TO THE FORMATION OF THE LIGAND-  
CC BINDING DOMAIN, WHILE THE BETA CHAINS CARRY THE KINASE DOMAIN (BY  
CC SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN  
CC RECEPTOR SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U72939; AAB17094.1; -  
CC HSSP; P06213; 1IRK.  
CC InterPro; IPR000494; EGFR\_L\_domain.  
CC InterPro; IPR000719; Euk\_pkinase.  
CC InterPro; IPR003961; FN\_III.  
CC InterPro; IPR002174; Furin-1ike.  
CC InterPro; IPR002011; RTKinasell.  
CC InterPro; IPR001245; Tyr\_pkinase.  
CC Pfam; PF000041; fn3; 1.  
CC Pfam; PF00069; pkinase; 1.  
CC Pfam; PF00757; Furin-1ike; 1.  
CC Pfam; PF01030; Recep\_L\_domain; 2.  
CC PRINTS; PR00109; TYRKINASE.  
CC ProDom; PD000001; Euk\_pkinase; 1.  
CC SMART; SM00060; FN3; 2.  
CC SMART; SM00261; FU; 1.  
CC SMART; SM00219; TYRK; 1.  
CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
CC PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
CC Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;  
KW Glycoprotein; ATP-binding; Phosphorylation; Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 753  
FT PROPEP 754 757  
FT CHAIN 758 1390  
FT DOMAIN 758 966  
FT TRANSMEM 967 987  
FT DOMAIN 988 1390  
FT NP\_BIND 1025 1301  
FT BINDING 1031 1039  
FT ACT\_SITE 1059 1059  
FT MOD\_RES 1162 1162  
FT CARBOHYD 1192 1192  
FT CARBOHYD 76 76  
FT CARBOHYD 196 196  
FT CARBOHYD 257 257  
FT CARBOHYD 322 322  
FT CARBOHYD 340 340  
FT CARBOHYD 439 439  
FT CARBOHYD 457 457  
FT CARBOHYD 492 492  
FT CARBOHYD 553 553  
FT CARBOHYD 765 765  
FT CARBOHYD 772 772  
FT CARBOHYD 793 793  
FT CARBOHYD 796 796  
FT CARBOHYD 871 871  
FT CARBOHYD 919 919  
FT CARBOHYD 933 933  
SO SEQUENCE 1390 AA; 156831 MW; 6BD2AA382EFD6442 CRC64;

Query Match 2.0%; Score 9; DB 1; length 1390;  
Best Local Similarity 100.0%; Pred. No. 1.7;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 191 SYGVVLMM 199  
| | | | | | | | | |  
Db 1224 SYGVVLMM 1232  
RESULT 8  
UL51 HCMVA STANDARD; PRT; 157 AA.  
ID UL51 HCMVA  
AC P16792;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last annotation update)  
DE Hypothetical protein UL51.  
GN UL51.  
OS Human cytomegalovirus (strain AD169).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Cytomegalovirus.  
OX NCBI\_TaxID=10360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90269039; PubMed=2161319;  
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,  
RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,  
RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.,  
RT "Analysis of the protein-coding content of the sequence of human  
RT cytomegalovirus strain AD169."  
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X17403; CAA35410.1; -  
CC PIR; S09814; S09814.  
CC InterPro; IPR005208; Herpes\_UL33.  
DR Pfam; PF03581; Herpes\_UL33; 1.  
DR KW Hypothetical protein.  
SO SEQUENCE 157 AA; 16968 MW; 5999035AC654517D CRC64;  
Query Match 1.8%; Score 8; DB 1; length 157;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 440 DDDGHEE 447  
| | | | | | | | | |  
Db 12 DDDGHEE 19  
RESULT 9  
2SS\_SOYBN STANDARD; PRT; 158 AA.  
ID 2SS\_SOYBN  
AC P19594;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 2S albumin precursor (GM2S-1) [Contains: Aspartic acid-rich peptide;  
DE 8 kDa methionine-rich protein (8 kDa MRP)].  
OS Glycine max (Soybean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eustoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
OX NCBI\_TaxID=3847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Hodgson 78; TISSUE=Cotyledon;  
RA Galvez A.F., Revilla M.J.R., de Lumen B.O.;  
RT "A novel methionine-rich protein from soybean cotyledon: cloning and

RT characterization of cDNA."  
RL (1n) Plant Gene Register FGR97-103.  
RN [2]  
RP SEQUENCE OF 22-64.  
RC TISSUE=Seed;  
RX MEDLINE=87280104; PubMed=3611081;  
RA Odani S., Koide T., Ono T.;  
RT "Amino acid sequence of a soybean (Glycine max) seed polypeptide  
RL having a poly(L-aspartic acid) structure."  
RN J. Biol. Chem. 262:10502-10505(1987).  
RP SEQUENCE OF 82-96.  
RA Revilla M.J., Galvez A.F., Krenz D.C., de Lumen B.O.;  
RT "An 8 kDa methionine-rich protein (MRP) from soybean (Glycine max)  
RL cotyledon: identification, purification and N-terminal sequence."  
RN J. Agric. Food Chem. 44:2930-2935(1996).  
CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.  
CC -1- SUBUNIT: THE PROTEIN CONSISTS OF TWO CHAINS LINKED BY 2 DISULFIDE  
CC BONDS.  
CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF005030; AAB71140.1; -.  
DR PIR; A28485; A28485.  
DR InterPro; IPR003612; AAI.  
DR InterPro; IPR001768; Try/amy1 inhbt.  
DR Pfam; PF00234; tryp\_alpha\_amy1; 1.  
DR SMART; SM00499; AAI; 1.  
KM Signal; Seed storage protein.  
FT SIGNAL 1 21  
FT CHAIN 22 64  
FT PROPEP 65 81  
FT CHAIN 82 158  
FT SITE 54 56  
FT DOMAIN 56 64  
FT DOMAIN 88 91  
SQ SEQUENCE 158 AA; 18460 MW; C1F42723B0F62D81 CRC64;  
Query Match  
Best Local Similarity 1.8%; Score 8; DB 1; Length 158;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 436 GDDDDDDDD 443  
DB 55 GDDDDDDDD 62  
RESULT 10  
REGO\_BP82 STANDARD; PRT; 229 AA.  
AC P13870;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Antiterminal protein Q.  
GN Q.  
OS Bacteriophage 82.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
OC Lambda-like viruses.  
OX NCBI\_TaxID=10705;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87308148; PubMed=3624233;  
RA Goliger J.A., Roberts J.W.;  
RT "Bacteriophage 82 gene Q and Q protein. Sequence, overproduction, and

RT activity as a transcription antiterminator in vitro."  
RL J. Biol. Chem. 262:11721-11725(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96196428; PubMed=8648624;  
RA Mahdi A.A., Sharples G.J., Mandal T.N., Lloyd R.G.;  
RT "Holiday junction resolvases encoded by homologous rusa genes in  
RL Escherichia coli K-12 and phage 82."  
RN J. Mol. Biol. 257:561-573(1996).  
CC -1- FUNCTION: POSITIVELY REGULATE EXPRESSION OF THE PHAGE LATE GENE  
CC OPERONS. BACTERIAL HOST RNA POLYMERASE MODIFIED BY ANTITERMINATION  
CC PROTEINS TRANSCRIBES THROUGH TERMINATION SITES THAT OTHERWISE  
CC PREVENT EXPRESSION OF THE REGULATED GENES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J02803; AAA32298.1; -.  
DR EMBL; X92588; CAA63332.1; -.  
DR PIR; A29791; PQBP82.  
KW Transcription regulation; Transcription termination; DNA-binding.  
FT DNA\_BIND 14 32  
FT SEQUENCE 229 AA; 26400 MW; E6BDABE818498667 CRC64;  
Query Match  
Best Local Similarity 1.8%; Score 8; DB 1; Length 229;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 290 TLRLKKL 297  
DB 142 TLRLKKL 149  
RESULT 11  
MCM1\_YEAST STANDARD; PRT; 286 AA.  
AC P11746;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Pheromone receptor transcription factor (GRM/PRTF protein).  
GN MCM1 OR FUN80 OR YMR043W OR YM9532.08.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90249735; PubMed=2159934;  
RA Ammerer G.;  
RT "Identification, purification, and cloning of a polypeptide  
RT (PRTF/GRM) that binds to mating-specific promoter elements in  
RT yeast."  
RN Genes Dev. 4:299-312(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89141759; PubMed=3066908;  
RA Passmore S., Maine G.T., Elble R., Christ C., Tye B.K.;  
RT "Saccharomyces cerevisiae protein involved in plasmid maintenance is  
RT necessary for mating of MAT alpha cells."  
RL J. Mol. Biol. 204:593-606(1988).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88030692; PubMed=3311883;  
RA Dubois E., Bercy J., Descamps F., Messenguy F.;  
RT "Characterization of two new genes essential for vegetative growth in  
RT Saccharomyces cerevisiae: nucleotide sequence determination and  
RT chromosome mapping.";

RL Gene 55:265-275(1987).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Odell C., Bowman S., Barrell B.G., Rajandream M.A.;  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF COMPLEX WITH MAT2.  
RX MEDLINE=98140702; PubMed=9490409;  
RA Tan S., Richmond T.J.;  
RT "Crystal structure of the yeast MATApha2/MCM1/DNA ternary complex.";  
RL Nature 391:660-666(1998).  
CC -1- FUNCTION: INTERACTS WITH THE ALPHA-2 REPRESSOR OR WITH THE  
CC ALPHA-1 ACTIVATOR THEREBY REGULATING THE EXPRESSION OF MATING-  
CC TYPE-SPECIFIC GENES.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION  
CC FACTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X52453; CAA36691.1; -.  
DR EMBL; X14187; CAA32389.1; -.  
DR EMBL; M17511; AAA34609.1; -.  
DR EMBL; Z48502; CAA88409.1; -.  
DR PIR; A34599; A34599.  
DR PDB; 1MM; 18-MAR-98.  
DR TRANSFAC; T00501; -.  
DR TRANSFAC; T00501; -.  
DR SGD; S0004646; MCM1.  
DR InterPro; IPR002100; TF\_MADSbox.  
DR Pfam; PF00319; SRF-TF; 1.  
DR PRINTS; PR00404; MADSDOMAIN.  
DR SMART; SM00432; MADS; 1.  
DR PROSITE; PS00350; MADS\_BOX\_1; 1.  
DR PROSITE; PS0066; MADS\_BOX\_2; 1.  
KW Transcription regulation; DNA-binding; Nuclear protein; 3D-structure.  
FT DOMAIN 18 72 MADS.  
FT DOMAIN 98 120 ASP/GLU-RICH (ACIDIC).  
FT CONFLICT 9 9 P -> S (IN REF. 3).  
FT CONFLICT 37 37 S -> F (IN REF. 3).  
FT CONFLICT 156 157 GA -> AR (IN REF. 3).  
FT CONFLICT 158 286 MISSING (IN REF. 3).  
FT CONFLICT 286 AA; 32802 MW; FD75B4BF549E9E3B CRC64;  
SEQUENCE

Query Match 1.8%; Score 8; DB 1; Length 286;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443  
Db 108 GDDDDDD 115

RESULT 12  
OSTP\_MOUSE  
ID OSTP\_MOUSE STANDARD; PRT; 294 AA.  
AC P10923; P19008;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-JUL-1989 (Rel. 11, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Osteopontin precursor (Bone sialoprotein 1) (Minopontin) (Early T  
DE lymphocyte activation 1 protein) (Secreted phosphoprotein 1) (SP-1)  
DE (2AR) (Calcium oxalate crystal growth inhibitor protein).  
GN SPPI OR SPP-1 OR OP OR ETA-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c; TISSUE=Liver;  
RX MEDLINE=90354433; PubMed=2387863;  
RA Miyazaki Y., Setoguchi M., Yoshida S.Y., Akizuki S., Yamamoto S.;  
RT "The mouse osteopontin gene. Expression in monocytic lineages and  
RT complete nucleotide sequence.";  
RL J. Biol. Chem. 265:14432-14438(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Macrophage;  
RX MEDLINE=89263742; PubMed=2726465;  
RA Miyazaki Y., Setoguchi M., Yoshida S., Higuchi Y., Akizuki S.,  
RA Yamamoto S.;  
RT "Nucleotide sequence of cDNA for mouse osteopontin-like protein.";  
RL Nucleic Acids Res. 17:3298-3298(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89255479; PubMed=2722855;  
RA Craig A.M., Smith J.H., Denhardt D.T.;  
RT "Osteopontin, a transformation-associated cell adhesion  
RT phosphoprotein, is induced by 12-O-tetradecanoylphorbol 13-acetate in  
RT mouse epidermis.";  
RL J. Biol. Chem. 264:9682-9689(1989).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89310352; PubMed=2787378;  
RA Patarca R., Freeman G.J., Singh R.P., Wei F.-Y., Durfee T.,  
RA Blattner F., Regnier D.C., Kozak C.A., Mock B.A., Morse H.C. III,  
RA Jereils T.R., Cantor H.;  
RT "Structural and functional studies of the early T lymphocyte  
RT activation 1 (Eta-1) gene. Definition of a novel T cell-dependent  
RT response associated with genetic resistance to bacterial infection.";  
RL J. Exp. Med. 170:145-161(1989).  
RN [5]  
RP SEQUENCE OF 17-37.  
RC TISSUE=Kidney;  
RX MEDLINE=93034441; PubMed=1414495;  
RA Worcester E.M., Blumenthal S.S., Beshensky A.M., Lewand D.L.;  
RT "The calcium oxalate crystal growth inhibitor protein produced by  
RT mouse kidney cortical cells in culture is osteopontin.";  
RL J. Bone Miner. Res. 7:1029-1036(1992).  
CC -1- FUNCTION: BINDS TIGHTLY TO HYDROXYAPATITE. APPEARS TO FORM AN  
CC INTEGRAL PART OF THE MINERALIZED MATRIX. PROBABLY IMPORTANT TO  
CC CELL-MATRIX INTERACTION.  
CC -1- SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3.  
CC -1- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES.  
CC -1- PTM: N- AND O-GLYCOSYLATED.  
CC -1- SIMILARITY: BELONGS TO THE OSTEOPONTIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X51834; CAA36132.1; -.  
DR EMBL; X13986; CAA32165.1; -.  
DR EMBL; J04806; AAA57265.1; -.  
DR EMBL; X16151; CAA34276.1; -.  
DR PIR; J10105; J10105.  
DR PIR; A33853; A33853.  
DR PIR; A37818; A37818.  
DR PIR; S12064; S12064.  
DR MGI; MGI:98389; Sppl.  
DR InterPro; IPR002038; Osteopontin.  
DR Pfam; PF00865; Osteopontin; 1.  
DR PRINTS; PR00216; OSTEOPONTIN.  
DR SMART; SM00017; OSTEO; 1.

```

DR PROSITE; PS00884; OSTEOPONTIN; 1.
KW Glycoprotein; Sialic acid; Bone; Cell adhesion; Phosphorylation;
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 294 OSTEOPONTIN.
FT SITE 144 146 CELL ATTACHMENT SITE.
FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 43 43 L -> P (IN REF. 4).
FT CONFLICT 99 99 E -> G (IN REF. 3).
FT CONFLICT 122 122 V -> F (IN REF. 2).
SQ SEQUENCE 294 AA; 32459 MW; 9D5F32D67ABC53EA CRC64;

Query Match
Best Local Similarity 1.8%; Score 8; DB 1; Length 294;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
Db 88 DDDDDDDG 95

RESULT 13
OSTP_RAT STANDARD; PRT; 317 AA.
AC P08721;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Osteopontin precursor (Bone sialoprotein 1) (Secreted phosphoprotein
DE 1) (SPP-1).
GN SPP1 OR SPP-1 OR 2B7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=93054745; PubMed=1429723;
RA Singh K., Mukherjee A.B., de Vouge M.W., Mukherjee B.B.;
RT "Differential processing of osteopontin transcripts in rat kidney-
RT and osteoblast-derived cell lines.";
RL J. Biol. Chem. 267:23847-23851(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87067405; PubMed=3024151;
RA Oldberg A., Franzen A., Heinegaard D.;
RT "Cloning and sequence analysis of rat bone sialoprotein (osteopontin)
RT cDNA reveals an Arg-Gly-Asp cell-binding sequence.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:8819-8823(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Smooth muscle;
RX MEDLINE=94013467; PubMed=8408622;
RA Giachelli C.M., Bae N., Almeida M., Denhardt D.T., Alpers C.E.,
RA Schwartz S.M.;
RT "Osteopontin is elevated during neointima formation in rat arteries
RT and is a novel component of human atherosclerotic plaques.";
RL J. Clin. Invest. 92:1686-1696(1993).
CC -1- FUNCTION: BINDS TIGHTLY TO HYDROXYAPATITE. APPEARS TO FORM AN
CC INTEGRAL PART OF THE MINERALIZED MATRIX. PROBABLY IMPORTANT TO
CC CELL-MATRIX INTERACTION.
CC -1- SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3.
CC -1- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES.
CC -1- SIMILARITY: BELONGS TO THE OSTEOPONTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M99252; AAA41765.1; -
DR EMBL; M14656; AAA41762.1; -.
DR PIR; A25917; A25917.
DR InterPro; IPR002038; Osteopontin.
DR Pfam; PF00865; Osteopontin; 1.
DR PRINTS; PR00216; OSTEOPONTIN.
DR SMART; SM00017; OSTEOPONTIN.
DR PROSITE; PS00884; OSTEOPONTIN; 1.
KW Glycoprotein; Sialic acid; Bone; Cell adhesion; Phosphorylation;
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 317 OSTEOPONTIN.
FT DOMAIN 86 96 POLY-ASP.
FT SITE 144 146 CELL ATTACHMENT SITE.
FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 8 8 F -> L (IN REF. 2).
SQ SEQUENCE 317 AA; 34963 MW; 73CB5C21FF62310 CRC64;

```

```

Query Match
Best Local Similarity 1.8%; Score 8; DB 1; Length 317;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
Db 88 DDDDDDDG 95

```

```

RESULT 14
RLA0_HALCU STANDARD; PRT; 352 AA.
ID RLA0_HALCU
AC P17006;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Acidic ribosomal protein P0 homolog (L10E).
GN RPLP0.
OS Halobacterium cutirubrum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=34001;
RX MEDLINE=89305527; PubMed=2743981;
RA Shimmitt L.C., Dennis P.P.;
RT "Characterization of the L11, L1, L10 and L12 equivalent ribosomal
RT protein gene cluster of the halophilic archaeobacterium Halobacterium
RT cutirubrum.";
RL EMBO J. 8:1225-1235(1989).
CC -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC OF E.COLI PROTEIN L10.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15078; CAA33180.1; -
DR PIR; S04120; R5H80C.
DR InterPro; IPR001790; Ribosomal L10.
DR Pfam; PF00466; Ribosomal_L10; 1.
KW Ribosomal protein.
SQ SEQUENCE 352 AA; 37199 MW; 2AE662E79A98D02D CRC64;

```

```

Query Match
Best Local Similarity 1.8%; Score 8; DB 1; Length 352;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



OY 437 DDDDDDDG 444  
|||||  
Db 333 DDDDDDDG 340

```
RESULT 15
RLAO_HALNI STANDARD; PRT; 352 AA.
ID RLAO_HALNI
AC P13553; Q9HQL4;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acidic ribosomal protein P0 homolog (L10E).
GN RPLP0 OR RPL10P OR VNG1104G.
OS Halobacterium sp. (strain NRC-1), and
OC Halobacterium halobium.
OC Archaea; Euryarchaeota; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091, 2242;
[1]
SEQUENCE FROM N.A.
SPECIES=H.halobium; STRAIN=S9;
MEDLINE=88329082; PubMed=2458258;
RC MEDLINE=88329082; PubMed=2458258;
RA Itoh T.;
RT "Complete nucleotide sequence of the ribosomal 'A' protein operon
RL from the archaeobacterium, Halobacterium halobium.";
RN Bur. J. Biochem. 176:297-303 (1988).
[2]
SEQUENCE FROM N.A.
STRAIN=NRC-1;
RC MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leitauer B., Keller K., Cruz R., Danson M.J., Hough D.W., Dale H.,
RA Maddocks D.G., Jablonksi P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Iisenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181 (2000).
CC -!- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC OF E.COLI PROTEIN L10.
CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X13008; CAA31431.1; -
CC EMBL; AE005042; AAG19499.1; -
DR EMBL; S01315; R5HSL0.
DR InterPro; IPR001790; Ribosomal_L10.
DR Pfam; PF00466; Ribosomal_L10; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 352 AA; 37213 MW; E2F088CF34FDB9BD CRC64;
```

Query Match 1.8%; Score 8; DB 1; Length 352;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
|||||  
Db 333,DDDDDDG 340

```
ID Y532 METJA STANDARD; PRT; 391 AA.
AC Q57952;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0532 precursor.
GN MJ0532.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
[1]
SEQUENCE FROM N.A.
STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RC MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Nguyen D.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Sadow P.W., Hanna M.C.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Borodovsky M.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073 (1996).
CC -!- SIMILARITY: BELONGS TO THE CHLP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67502; AAB98523.1; -
DR TIGR; MJ0532; -
DR InterPro; IPR000733; Flay_monooxygenase.
DR Pfam; PF01360; Monooxygenase; 1.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 391 HYPOTHETICAL PROTEIN MJ0532.
SQ SEQUENCE 391 AA; 43094 MW; A2F69AEEF8940EF2 CRC64;
```

Query Match 1.8%; Score 8; DB 1; Length 391;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 139 VVIADGV 146  
|||||  
Db 149 VVIADGV 156

```
RESULT 17
T2EA_YEAST STANDARD; PRT; 482 AA.
ID T2EA_YEAST
AC P36100;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transcription initiation factor IIE, alpha subunit (TFIIe-alpha)
DE (Transcription factor A large subunit) (Factor A 66 kDa subunit).
GN TPA1 OR YKL028W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 66-71; 102-108 AND 206-220.
STRAIN=BJ926;
RC MEDLINE=95050500; PubMed=7961670;
RA Feaver W.J., Henry N.L., Bushnell D.A., Sayre M.H., Brickner J.H.,
```

RA Gileadi O., Kornberg R.D.;  
RT "Yeast TFIIE. Cloning, expression, and homology to vertebrate  
RT proteins";  
RL J. Biol. Chem. 269:27549-27553 (1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Rieger M.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: RECRUITS TFIIF TO THE INITIATION COMPLEX AND STIMULATES  
CC THE RNA POLYMERASE II C-TERMINAL DOMAIN KINASE AND DNA-DEPENDENT  
CC ATPASE ACTIVITIES OF TFIIF. BOTH TFIIF AND TFIIE ARE REQUIRED FOR  
CC PROMOTER CLEARANCE BY RNA POLYMERASE (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Nucleus (Probable).  
CC -1- SIMILARITY: BELONGS TO THE TFIIE ALPHA SUBUNIT FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U12825; AAA62665.1; -  
CC EMBL; Z28028; CAA81863.1; -  
CC PIR; S37845; S37845.  
CC TRANSFAC; T02239; -  
CC SGD; S0001511; TPA1.  
CC InterPro; IPR002853; TFIIE\_alpha.  
CC Pfam; PF02002; TFIIE\_alpha; 1.  
CC SMART; SM00531; TFIIE; 1.  
CC Transcription regulation; Nuclear protein; Zinc-finger.  
CC ZN\_FING 124 152 C4-TYPE (POTENTIAL).  
CC DOMAIN 374 482 ASP/GLU-RICH (HIGHLY ACIDIC).  
CC FT 374 392 POLY-GLU.  
CC SQ SEQUENCE 482 AA; 54742 MW; 3E789DFCA247EF8A CRC64;  
Query Match 1.8%; Score 8; DB 1; Length 482;  
Best Local Similarity 100.0%; Pred. No. 6.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDDDD 443  
Db 464 GDDDDDDDD 471  
RESULT 18  
STK3\_HUMAN STANDARD; PRT; 487 AA.  
AC Q13043; Q15802; Q9NT24;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Serine/threonine protein kinase 4 (EC 2.7.1.37) (STE20-like kinase  
DE MST1) (MST-1) (Mammalian STE20-like protein kinase 1)  
DE (Serine/threonine protein kinase Krs-2).  
GN STK4 OR MST1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95394929; Pubmed=7665586;  
RA Creasy C.L., Chernoff J.;  
RT "Cloning and characterization of a human protein kinase with homology  
RT to Ste20";  
RL J. Biol. Chem. 270:21695-21700 (1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96413604; Pubmed=8816758;  
RA Taylor L.K., Wang H.C., Erikson R.L.;  
RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-

RT 2.";  
RT Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104 (1996).  
RN [3]  
RP SEQUENCE OF 1-435 FROM N.A.  
RA Laird G.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -1- ENZYME REGULATION: THE C-TERMINAL NON-CATALYTIC REGION INHIBITS  
CC THE KINASE ACTIVITY.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED.  
CC -1- PTM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC STE20 SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U18297; AAA83254.1; -  
CC EMBL; U60207; AAB17262.1; -  
CC EMBL; AL109839; CAB89421.1; -  
CC HSSP; P24941; IHCL.  
CC Genew; HGNC:11408; STK4.  
CC MIM; 604965; -  
CC InterPro; IPR000719; Euk\_pkinase.  
CC InterPro; IPR002290; Ser\_thr\_pkinase.  
CC Pfam; PF00069; pkinase; 1.  
CC PRINTS; PR00109; TYRKINASE.  
CC PRODOM; PD000001; Euk\_pkinase; 1.  
CC SMART; SM00220; S\_TKC; 1.  
CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE; PS00108; PROTEIN\_KINASE\_ST; FALSE\_NEG.  
CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
CC Transferrase; Serine/threonine-protein kinase; ATP-binding.  
CC DOMAIN 30 281  
CC NP\_BIND 36 44 ATP (BY SIMILARITY).  
CC BINDING 59 59 ATP (BY SIMILARITY).  
CC ACT\_SITE 149 149 BY SIMILARITY.  
CC DOMAIN 373 378 POLY-GLU.  
CC CONFLICT 222 222 P -> R (IN REF. 1).  
CC CONFLICT 312 312 V -> M (IN REF. 1).  
CC SQ SEQUENCE 487 AA; 55630 MW; 150758B5CF77D5C CRC64;  
Query Match 1.8%; Score 8; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 172 WMAPEVIQ 179  
Db 190 WMAPEVIQ 197  
RESULT 19  
STK3\_HUMAN STANDARD; PRT; 491 AA.  
AC Q13188; Q15801; Q15445;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Serine/threonine protein kinase 3 (EC 2.7.1.37) (STE20-like kinase  
DE MST2) (MST-2) (Mammalian STE20-like protein kinase 2)  
DE (Serine/threonine protein kinase Krs-1).  
GN STK3 OR MST2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96144292; PubMed=8566796;  
RA Creasy C.L., Chernoff J.;  
RT "Cloning and characterization of a member of the MST subfamily of  
RT Ste20-like kinases."  
RL Gene 167:303-306(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96413604; PubMed=8816758;  
RA Taylor L.K., Wang H.C., Erikson R.L.;  
RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-  
RT 2."  
RL Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).  
RN [3]  
RP SEQUENCE OF 96-203 FROM N.A.  
RX MEDLINE=94100173; PubMed=8274451;  
RA Schultz S.J., Nigg E.A.;  
RT "Identification of 21 novel human protein kinases, including 3 members  
RT of a family related to the cell cycle regulator nimA of Aspergillus  
RT nidulans."  
RL Cell Growth Differ. 4:821-830(1993).  
CC -1- FUNCTION: OXIDANT STRESS-ACTIVATED SERINE/THREONINE KINASE THAT  
CC MAY PLAY A ROLE IN THE RESPONSE TO ENVIRONMENTAL STRESS (BY  
CC SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN ADULT KIDNEY,  
CC SKELETAL AND PLACENTA TISSUES AND AT VERY LOW LEVELS IN ADULT  
CC HEART, LUNG AND BRAIN TISSUES.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC STE20 SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U26424; AAC50386.1; -  
DR EMBL; U60206; AAB17261.1; -  
DR EMBL; Z25422; CAA80909.1; -  
DR HSSP; P24941; 1HCL.  
DR Genew; HGNC:11406; STK3.  
DR MIM; 605030; -  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SMO0220; S\_TKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; FALSE\_NEG.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding.  
FT DOMAIN 27 278 PROTEIN KINASE.  
FT NP\_BIND 33 41 ATP (BY SIMILARITY).  
FT BINDING 56 56 ATP (BY SIMILARITY).  
FT ACT\_SITE 146 146 BY SIMILARITY.  
FT DOMAIN 308 314 POLY-GLU.  
FT DOMAIN 370 375 POLY-GLU.  
FT CONFLICT 96 98 WIV -> YLY (IN REF. 3).  
FT CONFLICT 121 121 D -> Y (IN REF. 3).  
FT CONFLICT 203 203 D -> G (IN REF. 3).  
FT CONFLICT 303 303 D -> E (IN REF. 2).  
FT CONFLICT 332 334 GEC -> ESV (IN REF. 2).  
SQ SEQUENCE 491 AA; 56261 MW; 9CA3B0644F3C14A9 CRC64;  
Query Match 1.8%; Score 8; DB 1; Length 491;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 WMAPEVIQ 179  
Db 187 WMAPEVIQ 194  
RESULT 20  
M3K7\_MOUSE . STANDARD; PRT; 579 AA.  
ID M3K7\_MOUSE  
AC Q62073;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)  
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-  
DE activated kinase 1).  
DE MAP3K7 OR TAK1.  
GN Mus musculus (Mouse).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96123277; PubMed=8533096;  
RA Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,  
RA Taniguchi T., Nishida E., Matsumoto K.;  
RT "Identification of a member of the MAPKK family as a potential  
RT mediator of TGF-beta signal transduction."  
RL Science 270:2008-2011(1995).  
CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.  
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B  
CC ACTIVATION.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC MAP KINASE KINASE SUBFAMILY.  
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D76446; BAA1184.1; -  
DR HSSP; P08631; 1AD5.  
DR MGD; MGI:1346877; Map3k7.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SMO0221; STYKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding.  
FT DOMAIN 8 16 POLY-SER.  
FT DOMAIN 36 291 PROTEIN KINASE.  
FT NP\_BIND 42 50 ATP (BY SIMILARITY).  
FT BINDING 63 63 ATP (BY SIMILARITY).  
FT ACT\_SITE 156 156 BY SIMILARITY.  
SQ SEQUENCE 579 AA; 64227 MW; 97C8F6F3C8E283EE CRC64;  
Query Match 1.8%; Score 8; DB 1; Length 579;  
Best Local Similarity 100.0%; Pred. No. 7.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 146 VLKICDFG 153  
Db 170 VLKICDFG 177  
RESULT 21

YM72\_YEAST  
ID YM72\_YEAST STANDARD; PRT; 590 AA.  
AC 005021;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Hypothetical 67.6 kDa protein in MRPL4-MTF1 intergenic region.  
GN YMR227C OR YMR959.09C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SS288C / AB972;  
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z49939; CAA90198.1; -  
DR TRANSFAC; T03088; -  
DR SGD; S0004840; YMR227C.  
KW Hypothetical protein; Coiled coil.  
FT DOMAIN 50 53 POLY-LYS.  
FT DOMAIN 200 203 POLY-GLU.  
FT DOMAIN 368 373 POLY-GLU.  
FT DOMAIN 413 421 POLY-ASP.  
FT DOMAIN 517 531 POLY-GLU.  
FT DOMAIN 427 549 COILED COIL (POTENTIAL).  
SQ SEQUENCE 590 AA; 67555 MW; C014E7419B0B1C61 CRC64;  
Query Match  
Best Local Similarity 1.8%; Score 8; DB 1; Length 590;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDDG 444  
DB 413 DDDDDDDG 420  
RESULT 22  
ID YM72\_HUMAN STANDARD; PRT; 606 AA.  
AC 043318; 043317; 043319;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.1 -)  
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-  
DE activated kinase 1).  
GN MAP3K7 OR TAK1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.  
RC TISSUE=Lung;  
RX MEDLINE=98153801; PubMed=9480845;  
RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.;  
RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an  
RT NF-kappa B-inducing kinase-independent mechanism.";  
RL Biochem. Biophys. Res. Commun. 243:545-549(1998).  
CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKs.  
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B  
CC ACTIVATION.

CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1A, 1B (SHOWN HERE) AND 1C, ARE  
CC PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC MAP KINASE KINASE KINASE SUBFAMILY.  
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, STRONG, TO C21ORF7.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AB009357; BAA25026.1; -  
DR EMBL; AB009356; BAA25025.1; -  
DR EMBL; AB009358; BAA25027.2; -  
DR HSSP; P08631; 1AD5.  
DR Genew; HGNC:6859; MAP3K7.  
DR MIM; 602614; -  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00221; STYKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
KW Alternative splicing.  
FT DOMAIN 8 14 POLY-SER.  
FT NP\_BIND 36 291 PROTEIN\_KINASE.  
FT BINDING 42 50 ATP (BY SIMILARITY).  
FT ACT\_SITE 63 63 ATP (BY SIMILARITY).  
FT ACT\_SITE 156 156 BY SIMILARITY.  
FT VARSPPLIC 404 430 MISSING. (IN ISOFORM 1A).  
FT VARSPPLIC 509 518 PLAPCNSKE -> ARTSCRTGPG (IN ISOFORM 1C).  
FT VARSPPLIC 519 606 MISSING (IN ISOFORM 1C).  
SQ SEQUENCE 606 AA; 67196 MW; 3D8F8147CD174013 CRC64;  
Query Match  
Best Local Similarity 1.8%; Score 8; DB 1; Length 606;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 146 VLKICDPG 153  
DB 170 VLKICDPG 177  
RESULT 23  
ID E2H2\_HUMAN STANDARD; PRT; 746 AA.  
AC Q15910; Q92857; Q15755;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Enhancer of zeste homolog 2 (ENX-1).  
GN E2H2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=97124843; PubMed=8954776;  
RA Chen H., Rossier C., Antonarakis S.E.;  
RT "Cloning of a human homolog of the Drosophila enhancer of zeste gene  
RT (E2H2) that maps to chromosome 21q22.2.";  
RL Genomics 38:30-37(1996).  
RN [2]  
RP SEQUENCE FROM N.A.



RX MEDLINE=97357309; PubMed=9214638;  
RA Laible G., Wolf A., Dorn R., Reuter G., Nislow C., Lebersorger A.,  
RA Popkin D., Pillus L., Jenuwein T.,  
RT "Mammalian homologues of the Polycomb-group gene Enhancer of zeste  
mediate gene silencing in Drosophila heterochromatin and at S.  
cerevisiae telomeres.";  
RL EMO. J. 16:3219-3232(1997).  
RN [3]  
RP SEQUENCE OF 134-746 FROM N.A.  
RX MEDLINE=96220494; PubMed=8649418;  
RA Hobert O., Jallat B., Ullrich A.,  
RT "Interaction of Vav with ENX-1, a putative transcriptional regulator  
of homeobox gene expression.";  
RL Mol. Cell. Biol. 16:3066-3073(1996).  
RN [4]  
RP ATRX BINDING.  
RX MEDLINE=98167853; PubMed=9499421;  
RA Cardoso C., Timsit S., Villard L., Khrestchatsky M., Fontes M.,  
RA Colleaux L.,  
RT "Specific interaction between the XNP/ATRX-X gene product and the SET  
domain of the human EZH2 protein.";  
Hum. Mol. Genet. 7:679-684(1998).  
CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF GENE TRANSCRIPTION  
AND CHROMATIN STRUCTURE.  
CC -1- SUBUNIT: BINDS ATRX VIA THE SET DOMAIN (PROBABLE).  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TISSUES.  
CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE EZ FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X95653; CAA64955.1; -  
DR EMBL; U61145; AAC51520.1; -  
DR EMBL; U52965; AAC50591.1; -  
DR TRANSFAC; T04888; -  
DR Genew; HGNC:3527; EZH2.  
DR MIM; 601573; -  
DR InterPro; IPR001005; Myb\_DNA\_binding.  
DR InterPro; IPR001214; SET.  
DR Pfam; PF00856; SET; 1.  
DR SMART; SM00395; SANT; 2.  
DR SMART; SM00317; SET; 1.  
DR PROSITE; PS50280; SET; 1.  
KW Transcription regulation; Nuclear protein; DNA-binding.  
FT DOMAIN 490 495 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
FT DOMAIN 523 605 CYS-RICH.  
FT DOMAIN 618 729 SET.  
FT DOMAIN 724 724 F -> L (IN REF. 1).  
FT CONFLICT 724 724 F -> V (IN REF. 1).  
FT CONFLICT 724 724 F -> V (IN REF. 1).  
SQ SEQUENCE 746 AA; 85363 MW; 1B5029EB9D509BES CRC64;  
Query Match 1.8%; Score 8; DB 1; Length 746;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 437 DDDDDDDG 444  
DB 183 DDDDDDDG 190  
RESULT 24  
EZHZ\_MOUSE STANDARD: PRT; 746 AA.  
ID EZHZ\_MOUSE  
AC 061188; Q9R090;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Enhancer of zeste homolog 2 (ENX-1).  
GN EZH2 OR ENX1H.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=97014262; PubMed=8861097;  
RA Hobert O., Sures I., Ciossek T., Fuchs M., Ullrich A.,  
RT "Isolation and developmental expression analysis of Enx-1, a novel  
mouse Polycomb group gene.";  
Mech. Dev. 55:171-184(1996).  
RN [2]  
RP SEQUENCE OF 134-497 FROM N.A.  
RC STRAIN=129/Sv;  
RX MEDLINE=99160476; PubMed=10051331;  
RA Laible G., Haynes A.R., Lebersorger A., O'Carroll D., Mattei M.G.,  
RA Denny P., Brown S.D., Jenuwein T.,  
RT "The murine polycomb-group genes *ezh1* and *ezh2* map close to *hox* gene  
clusters on mouse chromosomes 11 and 6.";  
Mamm. Genome 10:311-314(1999).  
RL Mamm. Genome 10:311-314(1999).  
CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF GENE TRANSCRIPTION  
AND CHROMATIN STRUCTURE.  
CC -1- SUBUNIT: BINDS ATRX VIA THE SET DOMAIN (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ENX-1A (SHOWN HERE) AND ENX-1B;  
ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED IN EARLY EMBRYOS. IN LATER  
EMBRYOGENESIS, EXPRESSION RESTRICTED TO CENTRAL AND PERIPHERAL  
NERVOUS SYSTEM, LIVER AND THYMUS. IN ADULT, HIGHEST EXPRESSION IN  
SPLEEN, TESTIS AND PLACENTA. LOWER LEVELS IN INTESTINE AND MUSCLE  
AND VERY LOW LEVELS IN BRAIN AND LIVER. NO EXPRESSION IN HEART,  
THYROID GLAND, LUNG AND KIDNEY.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN BOTH ADULT AND EMBRYO WITH  
HIGHEST LEVELS IN EARLY EMBRYOGENESIS.  
CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE EZ FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; U52951; AAC52655.1; -  
DR EMBL; AF104359; AAD54020.1; -  
DR MGD; MGI:107940; Ezh2.  
DR InterPro; IPR001005; Myb\_DNA\_binding.  
DR InterPro; IPR001214; SET.  
DR InterPro; IPR001214; SET.  
DR Pfam; PF00856; SET; 1.  
DR Pfam; PF00856; SET; 1.  
DR SMART; SM00395; SANT; 2.  
DR SMART; SM00317; SET; 1.  
DR PROSITE; PS50280; SET; 1.  
KW Transcription regulation; Nuclear protein; DNA-binding;  
KW Alternative splicing.  
FT DOMAIN 490 495 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
FT DOMAIN 523 605 CYS-RICH.  
FT DOMAIN 618 729 SET.  
FT DOMAIN 724 724 DGSNNHVVNYQPCDHPQPCDSSCPVIAQNFCEKFCQCSS  
FT VARSPLIC 511 553 EC -> G (IN ISOFORM ENX-1B).  
FT CONFLICT 159 161 MISSING (IN REF. 2).  
FT CONFLICT 746 AA; 85336 MW; 0435C021963ED24E CRC64;  
SQ SEQUENCE 746 AA; 85336 MW; 0435C021963ED24E CRC64;  
Query Match 1.8%; Score 8; DB 1; Length 746;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 437 DDDDDDDG 444

```
Db 183 DDDDDDDG 190
RESULT 25
CN3B_MOUSE
ID CN3B_MOUSE STANDARD; PRT; 799 AA.
AC Q61409;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE cGMP-inhibited 3',5'-cyclic phosphodiesterase B (EC 3.1.4.17) (Cyclic
GN GMP inhibited phosphodiesterase B) (CGI-PDE B) (CGIPDE1) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
SEQUENCE FROM N.A.
STRAIN=Swiss; TISSUE=Fat;
MEDLINE=97079687; PubMed=8921398;
RA Loebbert R.W., Winterpacht A., Selpel B., Zabel B.U.;
RT "Molecular cloning and chromosomal assignment of the human homologue
RT of the rat cGMP-inhibited phosphodiesterase 1 (PDE3A) -- a gene
RT involved in fat metabolism located at 11p15.1.";
RL Genomics 37:211-218(1996).
CC -1- FUNCTION: MAY PLAY A ROLE IN FAT METABOLISM.
CC -1- CATALYTIC ACTIVITY: Nucleoside 3',5'-cyclic phosphate + H(2)O =
CC nucleoside 5'-phosphate.
CC -1- ENZYME REGULATION: INHIBITED BY cGMP.
CC -1- SUBCELLULAR LOCATION: Membrane-bound (Potential).
CC -1- TISSUE SPECIFICITY: ABUNDANT IN ADIPOSE TISSUES.
CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X95521; CAA64775.1; -.
DR MGD; MGI:1333863; Pde3b.
DR InterPro; IPR003607; ME_P1ase_HDC.
DR InterPro; IPR002073; PDEase.
DR Pfam; PF00233; PDEase; 1.
DR SMART; SM00471; HDC; 1.
DR PROSITE; PS00126; PDEASE_1; 1.
KW Hydrolase; cGMP; Membrane.
FT NON_TER 1 1
FT DOMAIN 715 724 POLY-ASP.
FT DOMAIN 773 777 POLY-GLU.
FT NON_TER 799 799
SQ SEQUENCE 799 AA; 89811 MW; B76685895F1C8B3F CRC64;

Query Match 1.8%; Score 8; DB 1; Length 799;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
Db 716 DDDDDDDG 723

RESULT 26
YHT1_YEAST STANDARD; PRT; 840 AA.
AC P38635;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
```

```
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 95.1 kDa protein in ACT5-YCK1 intergenic region.
GN YHR131C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
SEQUENCE FROM N.A.
STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VII."
RL Science 265:2077-2082(1994).
CC -1- SIMILARITY: TO YEAST YNL144C.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U10398; AAB68414.1; -.
DR PIR; S48975; S48975.
DR HSSP; Q00963; IDRO.
DR SGD; S0001173; YHR131C.
DR InterPro; IPR001849; PH.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00233; PH; 1.
KW Hypothetical protein.
FT DOMAIN 312 325 POLY-ARG.
FT DOMAIN 788 840 ASP/GLU-RICH (HIGHLY ACIDIC).
FT NON_TER 840 840
SQ SEQUENCE 840 AA; 95058 MW; 3A86717D3332A0DF CRC64;

Query Match 1.8%; Score 8; DB 1; Length 840;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
Db 829 DDDDDDDG 836

RESULT 27
DD10_HUMAN STANDARD; PRT; 875 AA.
AC Q13206;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable ATP-dependent RNA helicase DDX10 (DEAD-box protein 10).
GN DDX10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=96301396; PubMed=8660968;
RA Savitsky K., Ziv Y., Bar-Shira A., Gilad S., Tagle D.A., Smith S.,
RA Uziel T., Sfez S., Nahmias J., Sartiell A., Eddy R.L., Shows T.B.,
RA Collins F.S., Shih Y., Rotman G.;
RT "A human gene (DDX10) encoding a putative DEAD-box RNA helicase at
11q22-q23.";
```

RL Genomics 33:199-206(1996).

RN [2]

RP SEQUENCE FROM N.A.

RA Arai Y., Kaneko Y., Kubo T., Arai K., Hosoda F., Ohki M.;

RT "Molecular analysis of the chromosomal breakpoints and identification of the repetitive sequences near the breakpoints of NUP98 in therapy-related leukemia with inv(11)(p15q22)."

RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE.

CC -1- TISSUE SPECIFICITY: HIGH IN TESTIS BUT WIDELY EXPRESSED.

CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DBP4/DDX10 SUBFAMILY.

CC -1- DATABASE: NAME=Atlas Genet. Cytoenet. Oncol. Haematol.;

CC WWW="http://www.infobiogen.fr/services/chromancer/Genes/DDX10.html".

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U28042; AAC50823.1; -.

DR EMBL; AB040537; BAB18536.1; -.

DR SWISS-2DPAGE; Q13206; HUMAN.

DR Genew; HGNC:2735; DDX10.

DR MIM; 601235; -.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR000629; DEAD\_box.

DR InterPro; IPR001650; Helicase\_C.

DR Pfam; PF00270; DEAD; 1.

DR Pfam; PF00271; Helicase\_C; 1.

DR SMART; SM00487; DEXDC; 1.

DR SMART; SM00490; HELIC; 1.

DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.

KW Helicase; ATP-binding; RNA-binding.

FT NP BIND 113 120 ATP (POTENTIAL).

FT SITE 222 225 DEAD BOX.

FT CONFLICT 647 647 A -> D (IN REF. 2).

FT CONFLICT 658 658 D -> E (IN REF. 2).

FT CONFLICT 661 661 N -> K (IN REF. 2).

SQ SEQUENCE 875 AA; 100815 MW; 4692EDA56AD945B5 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 875;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

437 DDDDDDDG 444

|||||

Db 782 DDDDDDDG 789

RESULT 28

CAFA\_HUMAN STANDARD; PRT; 938 AA.

ID Q13111; Q9UYI8; 40, Created

AC 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 41, Last annotation update)

DE Chromatin assembly factor 1 subunit A (CAF-1 subunit A) (Chromatin assembly factor 1 p150 subunit) (CAF-I 150 kDa subunit) (CAF-Ip150).

GN CHAF1A OR CAF1P150 OR CAF.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OC NCBI\_TaxID=9606;

OX [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Cervical adenocarcinoma;

RX MEDLINE=95323966; PubMed=7600578;

RA Kaufman P.D., Kobayashi R., Kessler N., Stillman B.;

RT "The p150 and p60 subunits of chromatin assembly factor I: a molecular

link between newly synthesized histones and DNA replication.";

RT Cell 81:1105-1114(1995).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).

RX MEDLINE=21147928; PubMed=11250073;

RA Dong H., Lin W., Zhang C.-K., Xiong H., Fu G., Jin W.-R., Chen R., Chen Z., Qi Z.-T., Huang G.M.;

RT "Genomic sequence and expression analyses of human chromatin assembly factor 1 p150 gene."

RT Gene 264:187-196(2001).

RL [3]

RP INTERACTION WITH PCNA, AND SUBCELLULAR LOCATION.

RX MEDLINE=99159824; PubMed=10052459;

RA Shibahara K., Stillman B.;

RT "Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin."

RL Cell 96:575-585(1999).

RN [4]

RP INTERACTION WITH PCNA.

RX MEDLINE=20115866; PubMed=10648606;

RA Moggs J.G., Grandi P., Quivy J.P., Jonsson Z.O., Hubscher U., Becker P.B., Almouzni G.;

RT "A CAF-1-PCNA-mediated chromatin assembly pathway triggered by sensing DNA damage."

RL Mol. Cell. Biol. 20:1206-1218(2000).

RN [5]

RP REVIEW.

RX MEDLINE=20351441; PubMed=10893180;

RA Ridgway P., Almouzni G.;

RT "CAF-1 and the inheritance of chromatin states: at the crossroads of DNA replication and repair."

RL J. Cell Sci. 113:2647-2658(2000).

CC -1- FUNCTION: COMPLEX THAT IS THOUGHT TO MEDIATE CHROMATIN ASSEMBLY IN DNA REPLICATION AND DNA REPAIR. ASSEMBLES HISTONE OCTAMERS ONTO REPLICATING DNA IN VITRO. CAF-1 PERFORMS THE FIRST STEP OF THE NUCLEOSOME ASSEMBLY PROCESS, BRINGING NEWLY SYNTHESIZED HISTONES H3 AND H4 TO REPLICATING DNA; HISTONES H2A/H2B CAN BIND TO THIS CHROMATIN PRECURSOR SUBSEQUENT TO DNA REPLICATION TO COMPLETE THE HISTONE OCTAMER. P150 BINDS TO HISTONES H3 AND H4. IT MAY PLAY A ROLE IN HETEROCHROMATIN MAINTENANCE IN PROLIFERATING CELLS BY BRINGING NEWLY SYNTHESIZED CBX PROTEINS TO HETEROCHROMATIC DNA REPLICATION FOCI (BY SIMILARITY).

CC -1- SUBUNIT: CAF-1 IS COMPOSED OF THREE SUBUNITS, P48, P60 AND P150. P150 BINDS DIRECTLY TO P60. ONLY MINOR AMOUNTS OF P48 ARE COMPLEXED WITH P60 AND P150 IN G1 PHASE. P150 BINDS DIRECTLY TO PCNA AND TO CBX1.

CC -1- SUBCELLULAR LOCATION: NUCLEAR; DNA REPLICATION FOCI.

CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; MAY BE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- DEVELOPMENTAL STAGE: ACTIVE COMPLEX IS FOUND IN G1, S AND G2 PHASES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U20979; AAA76736.1; -.

DR EMBL; AF190465; AAF04291.1; -.

DR Genew; HGNC:1910; CHAF1A.

DR MIM; 601246; -.

KW DNA replication; DNA repair; Cell cycle; Chaperone; Nuclear protein; Alternative splicing.

FT DOMAIN 1 31 BINDS TO PCNA.

FT DOMAIN 1 296 BINDS TO CBX1 CHROMO SHADOW DOMAIN.

FT DOMAIN 305 435 ARG/GLU/LYS-RICH.

FT DOMAIN 584 590 POLY-GLU.

FT DOMAIN 601 605 POLY-ASP.

FT DOMAIN 887 893 POLY-GLU.

FT DOMAIN 642 938 BINDS TO P60.



```

FT  VARSPLIC  692  731  CFELETPAQEEQTPKASKRERDEQIILPLHGNVNG
FT                                     -> HMVHPESRGDVCRITLRVSSPQSRYLNRNLNSCVKSTLSC
FT  VARSPLIC  732  938  MISSING (IN ISOFORM 2).
FT  VARSPLIC  754  926  MISSING (IN ISOFORM 3).
FT  CONFLICT  757  757  T -> S (IN REF. 2).
SQ  SEQUENCE  938 AA; 105222 MW; DBD01B9F1A0E28B7 CRC64;

Query Match      1.8%; Score 8; DB 1; Length 938;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  434 SEGDDDD 441
Db  598 SEGDDDD 605

RESULT 29
ROR2_HUMAN
ROR2_HUMAN  STANDARD;  PRT;  943 AA.
001974; Q9HAY7; Q9H8B1;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Tyrosine-protein kinase transmembrane receptor ROR2 precursor
(BC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 2).
GN  ROR2 OR NTRKR2.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX  NCBI_Taxid=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93100347; PubMed=1334494;
RA  Mastakowski P., Carroll R.D.;
RT  "A novel family of cell surface receptors with tyrosine kinase-like
RL  domain."
RT  J. Biol. Chem. 267:26181-26190(1992).
[2]
SEQUENCE OF 34-943 FROM N.A., AND VARIANT ILE-819.
MEDLINE=20164326; PubMed=10700182;
RA  Oldridge M., Fortuna A.M., Maringa M., Propping P., Mansour S.,
RA  Pollitt C., DeChiara T.M., Kimble R.B., Valenzuela D.M.,
RA  Yancopoulos G.D., Wilkie A.O.M.;
RT  "Dominant mutations in ROR2, encoding an orphan receptor tyrosine
RT  kinase, cause brachydactyly type B."
RL  Nat. Genet. 24:275-278(2000).
[3]
SEQUENCE OF 34-574 FROM N.A., AND VARIANT THR-245.
MEDLINE=20442029; PubMed=10986040;
RA  Schwabe G.C., Tinschert S., Buschow C., Meinecke P., Wolff G.,
RA  Gillesen-Kaeschbach G., Oldridge M., Wilkie A.O.M., Koemec R.,
RT  "Distinct mutations in the receptor tyrosine kinase gene ROR2 cause
RT  brachydactyly type B."
RL  Am. J. Hum. Genet. 67:822-831(2000).
[4]
VARIANTS RRS C-184; W-189; W-366 AND K-620.
MEDLINE=20392394; PubMed=10932186;
RA  Afzal A.R., Rajab A., Fenske C.D., Oldridge M., Elanko N.,
RA  Ternes-Pereira E., Tunesuez B., Murday V.A., Paton M.A.,
RA  Wilkie A.O.M., Jeffery S.;
RT  "Recessive Robinow syndrome, allelic to dominant brachydactyly type B,
RT  is caused by mutation of ROR2."
RL  Nat. Genet. 25:419-422(2000).
[5]
VARIANT RRS TYR-182.
MEDLINE=20392395; PubMed=10932187;
RA  van Bokhoven H., Celli J., Kayserili H., van Beusekom E., Balci S.,
RA  Brussel W., Skovby F., Kerr B., Percin E.F., Akarsu N., Brunner H.G.,
RT  "Mutation of the gene encoding the ROR2 tyrosine kinase causes
RT  autosomal recessive Robinow syndrome."
RL  Nat. Genet. 25:423-426(2000).

```

```

RN  [6]
RP  ERRATUM.
RA  van Bokhoven H., Celli J., Kayserili H., van Beusekom E., Balci S.,
RA  Brussel W., Skovby F., Kerr B., Percin E.F., Akarsu N., Brunner H.G.,
RL  Nat. Genet. 26:383-383(2000).
CC  -1- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHICH MAY BE INVOLVED
CC  IN THE EARLY FORMATION OF THE CHONDROCYTES. IT SEEMS TO BE
CC  REQUIRED FOR CARTILAGE AND GROWTH PLATE DEVELOPMENT.
CC  -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC  tyrosine phosphate.
CC  -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC  -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS DURING EARLY
CC  EMBRYONIC DEVELOPMENT. THE EXPRESSION LEVELS DROP STRONGLY AROUND
CC  DAY 16 AND THERE ARE ONLY VERY LOW LEVELS IN ADULT TISSUES.
CC  -1- DISEASE: DEFECTS IN ROR2 ARE A CAUSE OF BRACHYDACTYLY TYPE B
CC  (BDB). BDB IS AN AUTOSOMAL DOMINANT SKELETAL DISORDER
CC  CHARACTERIZED BY HYPOPLASIA/APLASIA OF DISTAL PHALANGES AND NAILS.
CC  IN BDB THE MIDDLE PHALANGES ARE SHORT BUT IN ADDITION THE TERMINAL
CC  PHALANGES ARE RUDIMENTARY OR ABSENT. BOTH FINGERS AND TOES ARE
CC  AFFECTED. THE THUMBS AND BIG TOES ARE USUALLY DEFORMED.
CC  -1- DISEASE: DEFECTS IN ROR2 ARE A CAUSE OF RECESSIVE ROBINOW SYNDROME
CC  (RRS). RRS IS AN AUTOSOMAL DISORDER CHARACTERIZED BY SKELETAL
CC  DYSPLASIA WITH GENERALIZED LIMB BONE SHORTENING, SEGMENTAL DEFECTS
CC  OF THE SPINE, BRACHYDACTYLY AND A DYSMORPHIC FACIAL APPEARANCE.
CC  -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
CC  SUBFAMILY.
CC  -1- SIMILARITY: CONTAINS 1 FRIZZLED (FZ) DOMAIN.
CC  -1- SIMILARITY: CONTAINS 1 IMMNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC  -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
DR  EMBL; M97639; AAA60276.1; -
DR  EMBL; AF294796; AAG01184.2; -
DR  EMBL; AF254747; AAG01184.2; JOINED.
DR  EMBL; AF254748; AAG01184.2; JOINED.
DR  EMBL; AF254749; AAG01184.2; JOINED.
DR  EMBL; AF254750; AAG01184.2; JOINED.
DR  EMBL; AF254751; AAG01184.2; JOINED.
DR  EMBL; AF254752; AAG01184.2; JOINED.
DR  EMBL; AF254753; AAG01184.2; JOINED.
DR  EMBL; AF279762; AAG33132.1; -
DR  EMBL; AF279755; AAG33132.1; JOINED.
DR  EMBL; AF279756; AAG33132.1; JOINED.
DR  EMBL; AF279757; AAG33132.1; JOINED.
DR  EMBL; AF279758; AAG33132.1; JOINED.
DR  EMBL; AF279759; AAG33132.1; JOINED.
DR  EMBL; AF279760; AAG33132.1; JOINED.
DR  EMBL; AF279761; AAG33132.1; JOINED.
DR  HSSP; P00747; ICRN.
DR  Genew; HGNC:10257; ROR2.
DR  MIM; 602337; -.
DR  MIM; 113000; -.
DR  MIM; 268310; -.
DR  InterPro; IPR000719; Euk_pkinase.
DR  InterPro; IPR000024; Fz_domain.
DR  InterPro; IPR003006; Ig_MHC.
DR  InterPro; IPR003598; Ig_c2.
DR  InterPro; IPR000001; Kringle.
DR  InterPro; IPR001245; Tyr_pkinase.
DR  Pfam; PF00047; Ig; 1.
DR  Pfam; PF00051; kringle; 1.
DR  Pfam; PF00069; pkinase; 1.
DR  Pfam; PF01392; Fz; 1.
DR  PRINTS; PR00018; KRINGLE.
DR  ProDom; PD000001; Euk_pkinase; 1.
DR  ProDom; PD000395; Kringle; 1.

```



DR SMART; SM00408; IGC2; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00219; TyKc; 1.  
DR PROSITE; PS50038; FZ; 1.  
DR PROSITE; PS00021; KRINGLE\_1; FALSE\_NEG.  
DR PROSITE; PS50070; KRINGLE\_2; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR Transferrase; Tyrosine-protein kinase; ATP-binding; Receptor;  
KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;  
KW Immunoglobulin domain; Developmental protein; Polymorphism;  
KW Disease mutation.  
FT SIGNAL 1 33  
FT CHAIN 34 943  
FT FT  
FT DOMAIN 34 403  
FT TRANSMEM 404 424  
FT FT 425 943  
FT DOMAIN 76 142  
FT DOMAIN 169 303  
FT DOMAIN 316 394  
FT DOMAIN 473 746  
FT DOMAIN 753 782  
FT DOMAIN 784 857  
FT DOMAIN 859 882  
FT NP BIND 479 487  
FT BINDING 507 507  
FT ACT SITE 615 615  
FT MOD RES 646 646  
FT DISTURB 83 135  
FT CARBOHYD 70 70  
FT CARBOHYD 188 188  
FT CARBOHYD 318 318  
FT CARBOHYD 182 182  
FT VARIANT 184 184  
FT VARIANT 189 189  
FT VARIANT 245 245  
FT VARIANT 366 366  
FT VARIANT 620 620  
FT VARIANT 819 819  
FT VARIANT 819 819  
FT SEQUENCE 943 AA; 104726 MW; DBAC1E4622B5ECA0 CRC64;  
Query Match 1.8%; Score 8; DB 1; Length 943;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 191 SYGVVLMW 198  
Db 677 SYGVVLMW 684  
RESULT 30  
ROR2\_MOUSE STANDARD; PRT; 944 AA.  
ID ROR2\_MOUSE  
AC 09Z138;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Tyrosine-protein kinase transmembrane receptor ROR2 precursor  
DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 2)  
DE (mROR2).  
GN ROR2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99248426; PubMed=10231392;  
RA Oishi I., Takeuchi S., Hashimoto R., Nagabukuro A., Ueda T., Liu Z.J.,  
RA Hata T., Akira S., Matsuda Y., Yamamura H., Otani H., Minami Y.,  
RT "Spatio-temporally regulated expression of receptor tyrosine kinases,  
RT mRor2, during mouse development: implications in development  
RT and function of the nervous system.";  
RL Genes Cells 4:41-56(1999).  
RN [2]  
RP CHARACTERIZATION.  
RX MEDLINE=20164325; PubMed=10700181;  
RA DeChiara T.M., Kimble R.B., Poueymirou W.T., Rojas J., Masiakowski P.,  
RA Valenzuela D.M., Yancopoulos G.D.,  
RT "Ror2, encoding a receptor-like tyrosine kinase, is required for  
RT cartilage and growth plate development.";  
RL Nat. Genet. 24:271-274(2000).  
CC -1- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHICH MAY BE INVOLVED  
CC IN THE EARLY FORMATION OF THE CHONDROCYTES. IT SEEMS TO BE  
CC REQUIRED FOR CARTILAGE AND GROWTH PLATE DEVELOPMENT.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR  
CC SUBFAMILY.  
CC -1- SIMILARITY: CONTAINS 1 FRIZZLED (FZ) DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AB010384; BAA75481.1; --  
CC HSSP; P00747; SHPG.  
CC MGI; MGI:1347521; Ror2.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR000024; Fz\_domain.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00047; Ig; 1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF01392; Fz; 1.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00408; IGC2; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00219; TyKc; 1.  
DR PROSITE; PS50038; FZ; 1.  
DR PROSITE; PS00021; KRINGLE\_1; FALSE\_NEG.  
DR PROSITE; PS50070; KRINGLE\_2; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Receptor;  
KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;  
KW Immunoglobulin domain; Developmental protein.  
FT SIGNAL 1 33  
FT CHAIN 34 944  
FT FT  
FT DOMAIN 34 403  
FT TRANSMEM 404 424  
FT DOMAIN 425 944  
CYTOPLASMIC (POTENTIAL).  
EXTRACELLULAR (POTENTIAL).  
POTENTIAL.

```

FT DOMAIN 76 142 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 169 303 FZ.
FT DOMAIN 316 394 KRINGLE.
FT DOMAIN 473 746 PROTEIN KINASE.
FT DOMAIN 753 782 SER/THR-RICH.
FT DOMAIN 784 857 PRO-RICH.
FT DOMAIN 859 882 SER/THR-RICH.
FT NP_BIND 479 487 ATP (BY SIMILARITY).
FT BINDING 507 507 ATP (BY SIMILARITY).
FT ACT_SITE 615 615 BY SIMILARITY.
FT MOD_RES 646 646 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT DISULFID 83 135 BY SIMILARITY.
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 944 AA; 105050 MW; CD2EEBC710387A56 CRC64;

```

```

Query Match
Best Local Similarity 1.8%; Score 8; DB 1; Length 944;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
191 SYGVVLWE 198
|||||
Db 677 SYGVVLWE 684

```

```

RESULT 31
VP41_YEAST
ID VP41_YEAST STANDARD; PRT; 992 AA.
AC P38959; Q12011; P87334;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Vacuolar assembly protein VPS41 (Vacuolar morphogenesis protein VAM2).
GN VPS41 OR VAM2 OR FET2 OR YDR080W OR D446 OR YD8554.13.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180-1A;
RX MEDLINE=97269041; PubMed=911041;
RA Nakamura N., Hirata A., Ohsumi Y., Wada Y.;
RT "Vam2/Vps41p and Vam6/Vps39p are components of a protein complex on
RT the vacuolar membranes and involved in the vacuolar assembly in the
RT yeast Saccharomyces cerevisiae."
RT J. Biol. Chem. 272:11344-11349(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=96093910; PubMed=7483840;
RA Coster F., Joniaux J.-L., Goffeau A.;
RT "Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open
RT reading frames, including TPS2, PPH3, RAD55, SED1, PDC2, AFR1, SSS1,
RT SLU7 and a tRNA for arginine."
RT Yeast 11:673-679(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Richards C., Harris D.E., Barrell B.G., Rajandream M.A.;
RN Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RP CHARACTERIZATION.
RX MEDLINE=97303186; PubMed=9159129;
RA Radisky D.C., Snyder W.B., Emr S.D., Kaplan J.;
RT "Characterization of VPS41, a gene required for vacuolar trafficking
RT and high-affinity iron transport in yeast."
RT Proc. Natl. Acad. Sci. U.S.A. 94:5662-5666(1997).
CC -1- FUNCTION: REQUIRED FOR VACUOLAR ASSEMBLY AND VACUOLAR TRAFFIC.
CC -1- SUBUNIT: INTERACTS WITH VPS39/VAM6.
CC -1- SIMILARITY: BELONGS TO THE VPS41 FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB000223; BAA19071.1; -.
DR EMBL; X82086; CAA57607.1; -.
DR EMBL; Z74376; CAA98899.1; -.
DR EMBL; Z46796; CAA86802.1; -.
DR PIR; S48767; S48767.
DR SGD; S0002487; VPS41.
DR InterPro; IPR00547; Clathrin_repeat.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 2.
DR Pfam; PF00637; Clathrin; 1.
DR SMART; SM00299; CLH; 1.
DR SMART; SM00320; WD40; 1.
FT DOMAIN 79 94 POLY-ASP.
FT DOMAIN 239 244 POLY-LYS.
FT CONFLICT 424 424 K -> M (IN REF. 2).
SQ SEQUENCE 992 AA; 113411 MW; E1E48B4D5A1A4005 CRC64;

```

```

Query Match
Best Local Similarity 1.8%; Score 8; DB 1; Length 992;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 436 GDDDDDDDD 443
|||||
Db 78 GDDDDDDDD 85

```

```

RESULT 32
IFH1_YEAST
ID IFH1_YEAST STANDARD; PRT; 1085 AA.
AC P39520;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE IFH1 protein (RRP3 protein).
GN IFH1 OR RRP3 OR YLR223C OR L8083.9.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FL100;
RX MEDLINE=95304839; PubMed=7785326;
RA Cherel I., Thuriaux P.;
RT "The IFH1 gene product interacts with a fork head protein in
RT Saccharomyces cerevisiae."
RT Yeast 11:261-270(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favell A., Fulton L., Gattung S., Greco T., Kirsten J.,
RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA Mardis E., Menezes S., Miller N., Nhan M., Trevaaskis E., Vignati D.,
RA Rifken L., Riles L., Taich A., Trevaaskis E., Vignati D.,
RA Wilcox L., Woldman P., Vaudin M., Wilson R., Waterston R.;
RN Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CONTROLS THE PRE-RRNA PROCESSING MACHINERY IN
CC CONJUNCTION WITH FHL1. COULD CONVERT FHL1 FROM A REPRESSOR
CC TO AN ACTIVATOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -----

```

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; Z29488; CAA82624.1; -  
DR EMBL; U19027; AAB67412.1; -  
DR PIR; S47477; S47477.  
DR SGD; S0004213; IFH1.  
KW Nuclear protein; Transcription regulation.  
FT DOMAIN 122 163 ASP/GLU-RICH (HIGHLY ACIDIC).  
SQ SEQUENCE 1085 AA; 122491 MW; BE1C7DEF06213FE0 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 1085;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
142 DDDDDDDG 149

RESULT 33  
YAB9\_YEAST STANDARD; PRT; 1131 AA.  
ID YAB9\_YEAST  
AC P31380;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 128.5 kDa helicase in ATSI-TPD3 intergenic region.  
GN YAL019W OR YAL001 OR FUN30.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=93209532; PubMed=8458570;  
RA Ouellette B.F.F., Clark M.W., Keng T., Storms R.K., Zhong W.W.,  
RA Zeng B., Fortin N., Delaney S., Barton A.B., Kaback D.B., Bussey H.,  
RT "Sequencing of chromosome I from Saccharomyces cerevisiae: analysis  
RT of a 32 kb region between the LTE1 and SPO7 genes.";  
RL Genome 36:32-42(1993).  
RN [2]  
RN SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=94193531; PubMed=8144453;  
RA Barton A.B., Kaback D.B.;  
RT "Molecular cloning of chromosome I DNA from Saccharomyces cerevisiae:  
RT analysis of the genes in the FUN38-MAK16-SPO7 region.";  
RL J. Bacteriol. 176:1872-1880(1994).  
RN [3]  
RN SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=92221690; PubMed=1561836;  
RA Clark M.W., Zhong W.W., Keng T., Storms R.K., Barton A.B.,  
RA Kaback D.B., Bussey H.;  
RT "Identification of a Saccharomyces cerevisiae homolog of the SNF2  
RT transcriptional regulator in the DNA sequence of an 8.6 kb region in  
RT the LTE1-CYS1 interval on the left arm of chromosome I.";  
RL Yeast 8:133-145(1992).  
RN [4]  
RN FUNCTION: MAY FUNCTION AS A NEGATIVE REGULATORY ELEMENT IN THE  
RN PROCESS OF UV DAMAGE REPAIR.  
RN [5]  
RN SUBCELLULAR LOCATION: Nuclear (Potential).  
RN [6]  
RN SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; L05146; AAC04938.1; -  
DR PIR; S22266; S22266.  
DR SGD; S0000017; FUN30.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR001650; Helicase\_C.  
DR InterPro; IPR000330; SNF2\_N.  
DR Pfam; PF00176; SNF2\_N; 1.  
DR Pfam; PF00271; helicase\_C; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR SMART; SM00490; HELICC; 1.  
KW Hypothetical protein; Nuclear protein; DNA-binding; Helicase;  
KW ATP-binding.  
FT NP\_BIND 597 604 ATP (POTENTIAL).  
FT SITE 703 706 DEGH BOX.  
FT SITE 703 706  
SQ SEQUENCE 1131 AA; 128506 MW; 7E61B9AB3A42ED2 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 1131;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDDD 443  
Db 524 GDDDDDDDD 531

RESULT 34  
YNF4\_YEAST STANDARD; PRT; 1165 AA.  
ID YNF4\_YEAST  
AC P53950;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Hypothetical 128.1 kDa protein in OMP2-MSG5 intergenic region.  
GN YNL054W OR N2467 OR YNL2467W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=S288C / FY1676;  
RX MEDLINE=96021608; PubMed=8533472;  
RA Berger P., Daignon F., Crozet M.;  
RT "The sequence of a 44 420 bp fragment located on the left arm of  
RT chromosome XIV from Saccharomyces cerevisiae.";  
RL Yeast 11:967-974(1995).  
RN [2]  
RN ERRATUM.  
RX MEDLINE=97060022; PubMed=8904343;  
RA Berger P., Daignon F., Crozet M.;  
RL Yeast 12:297-297(1996).  
RN [3]  
RN This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; U12141; AAA99658.1; -  
DR EMBL; Z71330; CAA95925.1; -  
DR SGD; S0004999; YNL054W.  
KW Hypothetical protein.  
SQ SEQUENCE 1165 AA; 128140 MW; 0414978AA3066CE7 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 1165;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437. DDDDDDDG 444  
 DB 1085 DDDDDDDG 1092

## RESULT 35

RNFH\_AZOVI

ID\_RNFH\_AZOVI STANDARD; PRT; 86 AA.

AC Q9F5Y0;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

GN Protein rnfH.

OS RNFH.

OC Azotobacter vinelandii.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Azotobacter.

OX NCBI\_TaxID=354;

RN [1]

SEQUENCE FROM N.A.

STRAIN=DT;

Rubio L.M., Rangaraj P., Roberts G.P., Ludden P.W.;

Cloning and mutational analysis of the Azotobacter vinelandii gene

encoding the dinitrogenase gamma subunit."

Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

-1- SIMILARITY: BELONGS TO THE UPF0125 (RNFH) FAMILY.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: AF302049; AAC29821.1;

DR InterPro; IPR005346; UPF0125.

DR Pfam; PF03658; UPF0125; 1.

FT DOMAIN 78 86 POLY-ASP.

SQ SEQUENCE 86 AA; 9632 MW; EBD84A7315BC25A9 CRC64;

OY 437 DDDDDDD 443

DB 78 DDDDDDD 84

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 86;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443

DB 78 DDDDDDD 84

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 86;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443

DB 78 DDDDDDD 84

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 86;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443

DB 78 DDDDDDD 84

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 86;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443

DB 78 DDDDDDD 84

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 86;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443

DB 78 DDDDDDD 84

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 86;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443

DB 78 DDDDDDD 84

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 86;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443

DB 78 DDDDDDD 84

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 86;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443

DB 78 DDDDDDD 84

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 86;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443

DB 78 DDDDDDD 84

Query Match

RX MEDLINE=86232584; PubMed=3520485;  
 RA Thomas C.M., Smith C.A.;  
 RT "The trfB region of broad host range plasmid RK2: the nucleotide  
 RT sequence reveals incC and key regulatory gene trfB/korA/korD as  
 RT overlapping genes.";  
 RT Nucleic Acids Res. 14:4453-4469(1986).  
 CC -1- FUNCTION: IN CONJUNCTION WITH KORB, INHIBITS THE TRANSCRIPTION OF  
 CC KILA, TRFA AND KORAB OPERONS. IN CONJUNCTION WITH KORC IS  
 CC RESPONSIBLE FOR THE NEGATIVE CONTROL OF KILC AND KILE OPERONS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X03962; CAA27596.1; -.

DR PIR; A03588; RGECKK.

KW Plasmid; Transcription regulation; Repressor; DNA-binding.

FT DNA\_BIND 37 56 H-T-H MOTIF (POTENTIAL).

SQ SEQUENCE 101 AA; 11306 MW; 25622FB8F98F21A5 CRC64;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 WEADAKK 249

DB 89 WEADAKK 95

Query Match

Best Local Similarity 1.5%; Score 7; DB 1; Length 101;

SEQUENCE FROM N.A.

STRAIN=ATCC 28583 / FL100;

RX MEDLINE=84236098; PubMed=6329732;

RA van Loon A.P.G.M., de Groot R.J., de Haan M., Dekker A., Grivell L.A.;

of the yeast ubiquinol-cytochrome c reductase complex 17 kDa protein (EC 1.10.2.2)

extremely high content of acidic amino acids.";

EMBO J. 3:1039-1043(1984).

RN [2]

SEQUENCE FROM N.A.

STRAIN=S288C / AB972;

RX MEDLINE=95400292; PubMed=7670463;

RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,

Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,

Yamazaki M., Tashiro H., Eki T.;

"Analysis of the nucleotide sequence of chromosome VI from

Saccharomyces cerevisiae.";

Nat. Genet. 10:261-268(1995).

RN [3]

SEQUENCE FROM N.A.

STRAIN=S288C / AB972;

RX MEDLINE=96287654; PubMed=8686381;

RA Eki T., Naitou M., Hagiwara H., Abe M., Ozawa M., Sasanuma S.-I.,

Sasanuma M., Tsuchiya Y., Shibata T., Watanabe K., Ono A.,

Yamazaki M.-A., Tashiro H., Hanaoka F., Murakami Y.;

"Map location and nucleotide sequence of korA, a key regulatory gene

of promiscuous plasmid RK2.";

Nucleic Acids Res. 11:7453-7469(1983).

RN [2]

SEQUENCE FROM N.A.

STRAIN=S288C / AB972;

RX MEDLINE=96287654; PubMed=8686381;

RA Eki T., Naitou M., Hagiwara H., Abe M., Ozawa M., Sasanuma S.-I.,

Sasanuma M., Tsuchiya Y., Shibata T., Watanabe K., Ono A.,

Yamazaki M.-A., Tashiro H., Hanaoka F., Murakami Y.;

"Map location and nucleotide sequence of korA, a key regulatory gene



```

RT "Fifteen open reading frames in a 30.8 kb region of the right arm of
RT chromosome VI from Saccharomyces cerevisiae."
RL Yeast 12:177-190(1996).
CC -1- FUNCTION: THIS IS A COMPONENT OF THE UBIQUINOL-CYTOCHROME C
CC REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH
CC IS PART OF THE MITOCHONDRIAL RESPIRATORY CHAIN. THIS PROTEIN MAY
CC MEDATE FORMATION OF THE COMPLEX BETWEEN CYTOCHROMES C AND C1.
CC -1- FUNCTION: THIS HIGHLY ACIDIC PROTEIN MAY MEDATE THE FORMATION OF
CC THE COMPLEX BETWEEN CYTOCHROMES C AND C1.
CC -1- CATALYTIC ACTIVITY: QH(2) + 2 ferricytochrome c = Q + 2
CC ferrocyclochrome c.
CC -1- SUBUNIT: FUNGI BCI COMPLEX CONTAINS 10 SUBUNITS; 3 RESPIRATORY
CC SUBUNITS, 2 CORE PROTEINS AND 5 LOW-MOLECULAR WEIGHT PROTEINS.
CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
CC -1- SIMILARITY: BELONGS TO THE UQCRH/QCR6 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X00551; CAA25220.1; -.
CC EMBL; D50617; BAA09272.1; -.
CC PIR; A00120; RDBYUC.
CC SGD; S0001929; QCR6.
CC Mitochondrion; Inner membrane; Electron transport; Respiratory chain;
CC Oxidoreductase.
CC FT DOMAIN 26 80 ASP/GLU-RICH (ACIDIC).
CC FT CONFLICT 2 2 G -> D (IN REF. 1).
CC SQ SEQUENCE 147 AA; 17257 MW; 1488B9C0EB57BE8F CRC64;
CC
CC Query Match 1.5%; Score 7; DB 1; Length 147;
CC Best Local Similarity 100.0%; Pred. No. 22;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDD 443
Db 59 DDDDDDD 65

```

```

RL Science 269:496-512(1995).
RN [2]
RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=99182411; PubMed=1008086;
RA Hennig M., Dale G.E., D'Arcy A., Daniel F., Fischer S., Gray C.P.,
RA Jolidon S., Mueller F., Page M.G.P., Pattison P., Oefner C.;
RT "The structure and function of the 6-hydroxymethyl-7,8-dihydropterin
RT pyrophosphokinase from Haemophilus influenzae."
RL J. Mol. Biol. 287:211-219(1999).
CC -1- CATALYTIC ACTIVITY: ATP + 2-amino-4-hydroxy-6-hydroxymethyl-7,8-
CC dihydropteridine = AMP + 2-amino-7,8-dihydro-4-hydroxy-6-
CC (dihydrooxymethyl)pteridine.
CC -1- PATHWAY: Dihydrofolate biosynthesis; first step.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO THE HPPK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32691; AAC21742.1; -.
CC PDB; 1CBK; 01-MAR-00.
CC TIGR; HI0064; -.
CC InterPro; IPR000550; Hppk.
CC Pfam; PF01288; HPPK; 1.
CC PROSITE; PS00794; HPPK; 1.
CC KW Transferase; Folate biosynthesis; 3D-structure; Complete proteome.
CC SQ SEQUENCE 160 AA; 18299 MW; 136CD15F8844FDD CRC64;
CC
CC Query Match 1.5%; Score 7; DB 1; Length 160;
CC Best Local Similarity 100.0%; Pred. No. 24;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 222 NERLTTIP 228
Db 108 NERLTTIP 114

```

```

RESULT 38
HPPK_HAEIN STANDARD; PRT; 160 AA.
ID HPPK_HAEIN
AC P43777;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 2-amino-4-hydroxy-6-hydroxymethyl-dihydropteridine pyrophosphokinase
DE (EC 2.7.6.3) (7,8-dihydro-6-hydroxymethyl-pterin-pyrophosphokinase)
DE (HPPK) (6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase) (PPPK).
DE FOLK OR HI0064.
GN Haemophilus influenzae.
OS Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OC NCBI_TaxID=727;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Rd / KW20 / ATCC 51907;
RC MEDLINE=95350630; PubMed=7542800;
RX Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RX Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RX McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RX Scott J.D., Shirley R., Liu L.-L., Glodek A., Kelley J.M.,
RX Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RX Uutterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RX Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RX Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RX Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."

```

```

RESULT 39
UBC2_YEAST STANDARD; PRT; 172 AA.
ID UBC2_YEAST
AC P06104;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-20 kDa (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
DE UBC2 OR RAD6 OR YGL058W.
GN Saccharomyces cerevisiae (Baker's yeast).
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OC NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=85113143; PubMed=3881753;
RX Reynolds P., Weber S., Prakash L.;
RA "RAD6 gene of Saccharomyces cerevisiae encodes a protein containing a
RA tract of 13 consecutive aspartates."
RL Proc. Natl. Acad. Sci. U.S.A. 82:168-172(1985).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=S288C;
RC MEDLINE=97377993; PubMed=9234674;
RX Feuerhahn M., de Montigny J., Potier S., Souciet J.-L.;
RA "The characterization of two new clusters of duplicated genes
RA suggests a 'lego' organization of the yeast Saccharomyces cerevisiae
RA chromosomes."
RL Yeast 13:861-869(1997).
RN [3]

```

RP SEQUENCE OF 77-91, AND FUNCTION.  
RX MEDLINE=87315384; PubMed=3306404;  
RA Jentack S., McGrath J.P., Varshavsky A.;  
RT "The yeast DNA repair gene RAD6 encodes a ubiquitin-conjugating  
enzyme.";  
RL Nature 329:131-134(1987).  
RN [4]  
RP MUTAGENESIS OF CYS-88.  
RX MEDLINE=90207263; PubMed=2157209;  
RA Sung P., Prakash S., Prakash L.;  
RT "Mutation of cysteine-88 in the Saccharomyces cerevisiae RAD6 protein  
abolishes its ubiquitin-conjugating activity and its various  
biological functions.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:2695-2699(1990).  
RN [5]  
RP REQUIREMENT FOR E3.  
RX MEDLINE=91293094; PubMed=2065660;  
RA Sung P., Berleth E., Pickart C., Prakash S., Prakash L.;  
RT "Yeast RAD6 encoded ubiquitin conjugating enzyme mediates protein  
degradation dependent on the N-end-recognizing E3 enzyme.";  
RL EMBO J. 10:2187-2193(1991).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
RX MEDLINE=98165808; PubMed=9497353;  
RA Worthylake D.K., Prakash S., Prakash L., Hill C.P.;  
RT "Crystal structure of the Saccharomyces cerevisiae ubiquitin-  
conjugating enzyme Rad6 at 2.6-A resolution.";  
RL J. Biol. Chem. 273:6271-6276(1998).  
RN [7]  
RP FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO  
OTHER PROTEINS. UBC2 IS ACTIVE ON HISTONES. IT IS REQUIRED FOR  
POSTREPLICATION REPAIR OF UV-DAMAGED DNA AND SPOULATION. UBC2  
CC MEDATES E3-DEPENDENT UBC ACTIVITY.  
CC CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
CC phosphate + protein N-ubiquityllysine.  
CC -1- PATHWAY: Ubiquitin conjugation; second step.  
CC -1- SUBUNIT: Interacts with Rad18.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- DOMAIN: THE ACIDIC-TAIL DOMAIN OF RAD6 IS IMPORTANT FOR  
CC POLYUBQUITINATION OF HISTONES, AS WELL AS FOR SPOULATION.  
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR  
CC UBIQUITIN-THIOLESTER FORMATION.  
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; K02962; AAA34952.1; -  
DR EMBL; Z72580; CAA96761.1; -  
DR PIR; A21906; A21906.  
DR PDB; 1AYZ; 26-AUG-98.  
DR SGD; S0003026; RAD6.  
DR InterPro; IPR000608; UBC\_conjugat.  
DR Pfam; PF00179; UQ\_con; 1.  
DR ProDom; PD000461; UBC\_conjugat; 1.  
DR SMART; SM00212; UBCc; 1.  
DR PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
DR PROSITE; PS50127; UBIQUITIN\_CONJUGAT\_2; 1.  
KW Ubl conjugation pathway; Ligase; DNA repair; Sporulation;  
KW Nuclear protein; Multigene family; 3D-structure.  
FT DOMAIN 150 172 ASP/GLU-RICH (ACIDIC TAIL).  
FT MOD RES 1 1 BLOCKED.  
FT BINDING 88 88 UBIQUITIN.  
FT MUTAGEN 88 88 C->A,V: LOSS OF ACTIVITY.  
SQ SEQUENCE 172 AA; 19705 MW; 5F568DC28ABBD60F CRC64;

Query Match 1.5%; Score 7; DB 1; Length 172;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443  
Db 157 DDDDDDD 163  
RESULT 40  
NU6M\_BRABR STANDARD; PRT; 173 AA.  
AC P43194;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).  
GN MTND6 OR ND6 OR NADH6.  
OS Brachyramphus brevirostris (Kitlitz murrelet).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Charadriiformes; Alcidae;  
OC Brachyramphus.  
OX NCBI\_TaxID=28693;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94363783; PubMed=8082208;  
RA Moun T., Willassen N.P., Johansen S.;  
RT "Intragenic rearrangements in the mitochondrial NADH dehydrogenase  
subunit 6 gene of vertebrates.";  
RL Curr. Genet. 25:554-557(1994).  
RN [2]  
RP FUNCTION: NADH + ubiquinone = NAD(+) + ubiquinol.  
CC CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X73922; CAA52127.1; -  
DR InterPro; IPR001457; Oxidored\_q3.  
DR Pfam; PF00499; oxidored\_q3; 1.  
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
SQ SEQUENCE 173 AA; 18334 MW; 307D4AD6AC8C3419 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 173;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SLGASPV 9  
Db 45 SLGASPV 51  
RESULT 41  
NU6M\_BRAMA STANDARD; PRT; 173 AA.  
AC P43195;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).  
GN MTND6 OR ND6 OR NADH6.  
OS Brachyramphus marmoratus (Marbled murrelet).  
OC Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Charadriiformes; Alcidae;  
OC Brachyramphus.  
OX NCBI\_TaxID=28694;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94363783; PubMed=8082208;  
RA Moun T., Willassen N.P., Johansen S.;  
RT "Intragenic rearrangements in the mitochondrial NADH dehydrogenase

RT subunit 6 gene of vertebrates." ;  
RL Curr. Genet. 25:554-557(1994) .  
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X73923; CAA52128.1; -  
DR InterPro; IPR001457; Oxidored\_q3.  
DR Pfam; PF00499; oxidored\_q3; 1.  
KM Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
SQ SEQUENCE 173 AA; 18444 MW; F4065804ADF85499 CRC64;  
  
Query Match 1.5%; Score 7; DB 1; Length 173;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 3 SLGASFV 9  
Db 45 SLGASFV 51  
  
RESULT 42  
VIRR\_AGRV STANDARD; PRT; 175 AA.  
ID VIRR\_AGRV 052278;  
AC 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Virr protein.  
GN VIRR.  
OS Agrobacterium tumefaciens.  
OG Plasmid\*PTiA6NC.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=358;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A6NC;  
RA Zhu J., Winans S.C.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE UPF0066 (VIRR) FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF242881; AAB91563.1; -  
DR InterPro; IPR001378; UPF0066.  
DR Pfam; PF01980; UPF0066; 1.  
DR ProDom; PD006705; UPF0066; 1.  
DR TIGRFAMs; TIGR00104; UPF0066; 1.  
DR PROSITE; PS01318; UPF0066; 1.  
KM Plasmid.  
SQ SEQUENCE 175 AA; 18820 MW; 18AF20BB611B087F CRC64;  
  
Query Match 1.5%; Score 7; DB 1; Length 175;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 392 AGAVMHS 398  
Db 9 AGAVMHS 15

RESULT 43  
NO29\_XENLA STANDARD; PRT; 183 AA.  
ID NO29\_XENLA 042584;  
AC 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Nucleoplasmin-like protein NO29 (NOVA).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=97470976; PubMed=9326619;  
RA Zirmes R.R.F., Schmidt-Zachmann M.S., Franke W.W.;  
RT "Identification of a small, very acidic constitutive nucleolar protein  
RT (NO29) as a member of the nucleoplasmin family.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:11387-11392(1997).  
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.  
CC -1- SIMILARITY: BELONGS TO THE NUCLEOPLASMIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z85983; CAB06652.1; -  
DR InterPro; IPR004301; Nucleoplasmin.  
DR Pfam; PF03066; Nucleoplasmin; 1.  
KM Nuclear protein.  
FT DOMAIN 127 165 ASP/GLU-RICH (HIGHLY ACIDIC).  
FT DOMAIN 135 145 POLY-GLU.  
FT DOMAIN 146 162 POLY-ASP.  
SQ SEQUENCE 183 AA; 20121 MW; C1F5DE503FCE7E9F CRC64;  
  
Query Match 1.5%; Score 7; DB 1; Length 183;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 437 DDDDDDD 443  
Db 146 DDDDDDD 152  
  
RESULT 44  
HMG1\_ONCMY STANDARD; PRT; 204 AA.  
ID HMG1\_ONCMY P07746; Q91200;  
AC 01-AUG-1988 (Rel. 08, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE High mobility group-T protein (HMG-T) (HMG-T1) (HMG-1).  
OS Onchorynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Onchorynchus.  
OX NCBI\_TaxID=8022;  
RN [1].  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85269614; PubMed=4022777;  
RA Pentecost B.T., Wright J.M., Dixon G.H.;  
RT "Isolation and sequence of cDNA clones coding for a member of the  
RT family of high mobility group proteins (HMG-T) in trout and analysis  
RT of HMG-T-mRNA's in trout tissues.";  
RL Nucleic Acids Res. 13:4871-4888(1985).  
RN [2]

```

RP REVISIONS.
RA Wright J.;
RL Submitted (MAY-1986) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=95045507; PubMed=7957172;
RA Stros M., Mishikawa S., Dixon G.H.;
RT "CDNA sequence and structure of a gene encoding trout testis high-
mobility-group-1 protein.";
RL Eur. J. Biochem. 225:581-591(1994).
CC -1- FUNCTION: BINDS PREFERENTIALLY SINGLE-STRANDED DNA AND UNWINDS
CC DOUBLE STRANDED DNA.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE HMG1/HMG2 PROTEIN FAMILY.
CC -1- SIMILARITY: CONTAINS 2 HMG BOXES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X02666; CAA26500.1; -
CC EMBL; L32859; AAA58771.1; -
CC PIR; A24019; A24019.
CC HSSP; P07155; 1AAB.
CC InterPro; IPR000910; HMG_12_box.
CC InterPro; IPR000135; Highmobility_12.
CC Pfam; PF00505; HMG_box; 2.
CC PRINTS; PR00886; HGMHMOBLTY12.
CC SMART; SM00398; HMG; 2.
CC PROSITE; PS00353; HMG1_2; 1.
CC Nuclear protein; Chromosomal
CC DNA_BIND 8 78 protein; DNA-binding; Repeat.
CC FT DNA_BIND 8 78 HMG BOX 1.
CC FT DNA_BIND 94 162 HMG BOX 2.
CC FT DOMAIN 181 204 ASP/GLU-RICH (ACIDIC).
CC FT CONFLICT 22 22 R -> C (IN REF. 3).
CC FT CONFLICT 149 149 R -> K (IN REF. 3).
CC SQ SEQUENCE 204 AA; 23572 MW; 6A858A14AD669A0 CRC64;

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 204;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

437 DDDDDDD 443
|||||
181 DDDDDDD 187

RESULT 45
HIT_MOUSE
ID HIT_MOUSE STANDARD; PRT; 207 AA.
AC Q07133;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Histone H1c (Testicular H1 histone).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Leukocyte;
RX MEDLINE=94060108; PubMed=8241275;
RA Dracut B., Bode C., Doenecke D.;
RT "Structure and expression of the mouse testicular H1 histone gene
(h1c).";
RL Biochim. Biophys. Acta 1216:311-313(1993).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Testis;
RA van Wert J., Wright J., Wolfe S.A., Grimes S.R.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: HISTONES H1 ARE NECESSARY FOR THE CONDENSATION OF
CC NUCLEOSOME CHAINS INTO HIGHER ORDER STRUCTURES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DEVELOPMENTAL STAGE: THIS HISTONE IS A TESTIS-SPECIFIC H1 VARIANT
CC THAT APPEARS DURING MEIOSIS IN SPERMATOGENESIS.
CC -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L28753; AAB38417.1; -
CC EMBL; X72805; CAA51325.1; -
CC EMBL; U06232; AAA18359.1; -
CC PIR; S33226; S33226.
CC PIR; S43434; S43434.
CC HSSP; P08287; 1GHC.
CC MGD; MGI:107502; H1f3.
CC InterPro; IPR001386; Histone_H1/H5.
CC Pfam; PF00538; linker histone; 1.
CC ProDom; PD000373; Linkerhist_N; 1.
CC SMART; SM00526; H15; 1.
CC Chromosomal protein; Nuclear protein; DNA-binding; Multigene family;
CC Acetylation; Spermatogenesis; Testis.
CC FT INIT MET 0 0 BY SIMILARITY.
CC FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
CC FT DOMAIN 37 110 GLOBULAR.
CC FT CONFLICT 163 163 R -> G (IN REF. 1).
CC SQ SEQUENCE 207 AA; 21508 MW; A3C847CA97976C44 CRC64;

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 207;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 SKRRGKK 421
|||||
DB 21 SKRRGKK 27

RESULT 46
EF1B_ORYSA
ID EF1B_ORYSA STANDARD; PRT; 228 AA.
AC Q40680;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Elongation factor 1-beta (Ef-1-beta).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. HayaYuki; TISSUE=Anther;
RX MEDLINE=94139885; PubMed=8307147;
RA Matsumoto S., Terui Y., Shixiong X., Taira H., Ejiri S.I.;
RT "Cloning and characterization of the cDNA encoding rice elongation
factor 1 beta.";
RL FEBS Lett. 338:103-106(1994).
CC -1- FUNCTION: EF-1-BETA AND EF-1-BETA' STIMULATE THE EXCHANGE OF
CC GDP BOUND TO EF-1-ALPHA TO GTP.
CC -1- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,
CC BETA', AND GAMMA.

```



# ----- -1- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY. -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D23674; BAA04903.1; -  
DR InterPro; IPR001326; EF1\_BD.  
DR Pfam; PF00736; EF1BD; 1.  
DR PROSITE; PS00824; EF1BD\_1; 1.  
DR PROSITE; PS00825; EF1BD\_2; 1.  
KW Elongation factor; Protein biosynthesis.  
FT INIT MET 0 BY SIMILARITY.  
SQ SEQUENCE 228 AA; 24731 MW; D3CE96C9A7B4EF8D CRC64;

Query Match 1.5%; Score 7; DB 1; Length 228;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 100 DDDDDDD 106

## RESULT 47

EF1B\_BETVU STANDARD; PRT; 230 AA.  
ID EF1B\_BETVU  
AC 081918;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Elongation factor 1-beta (EF-1-beta).  
OS Beta vulgaris (Sugar beet).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.  
OX NCBI\_TaxID=3555;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. INZUCHTLINIE KMS VV-D/ZR5; TISSUE=leaf;  
RA Viereck R.;  
RT "Nucleotide sequence of the elongation factor 1-beta from Beta  
vulgaris.";  
Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: EF-1-BETA AND EF-1-BETA' STIMULATE THE EXCHANGE OF  
GDP BOUND TO EF-1-ALPHA TO GTP.  
-1- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,  
BETA', AND GAMMA.

-1- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.  
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; Z97067; CAB09803.1; -  
DR InterPro; IPR001326; EF1\_BD.  
DR Pfam; PF00736; EF1BD; 1.  
DR PROSITE; PS00824; EF1BD\_1; 1.  
DR PROSITE; PS00825; EF1BD\_2; 1.  
KW Elongation factor; Protein biosynthesis.  
FT INIT MET 0 BY SIMILARITY.  
SQ SEQUENCE 230 AA; 24545 MW; EBF120B8FEC41EAO CRC64;

Query Match 1.5%; Score 7; DB 1; Length 230;  
Best Local Similarity 100.0%; Pred. No. 34;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 101 DDDDDDD 107

## RESULT 48

ICP3\_HSVIN STANDARD; PRT; 245 AA.  
ID ICP3\_HSVIN  
AC P37319;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE Infected cell protein ICP34.5 (Neurovirulence factor ICP34.5).  
GN ICP34.5.  
OS Herpes simplex virus (type 1 / strain MGH-10).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=37107;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90156494; PubMed=2154589;  
RA Chou J., Roizman B.;  
RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted  
RT repeats, is conserved in several limited-passage isolates but not in  
RT strain 17syn+.";  
RL J. Virol. 64:1014-1020 (1990).  
CC -1- FUNCTION: NOT KNOWN; HAS A PROFOUND EFFECT ON THE GROWTH OF THE  
CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.  
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.  
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M33700; AAA45791.1; -  
KW Repeat; Late protein.  
FT DOMAIN 155 172 6 X 3 AA TANDEM REPEATS OF A-T-P.  
SQ SEQUENCE 245 AA; 25981 MW; 55ADA50AB1B8CA11 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 245;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 69 DDDDDDD 75

ICP3\_HSV11 STANDARD; PRT; 248 AA.  
ID ICP3\_HSV11  
AC P36313;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Infected cell protein ICP34.5 (Neurovirulence factor ICP34.5).  
GN ICP34.5.  
OS Herpes simplex virus (type 1 / strain 17).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=10299;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88274327; PubMed=2839594;  
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,  
McNab D., Perry L.J., Scott J.E., Taylor P.;

RT "The complete DNA sequence of the long unique region in the genome of  
RT herpes simplex virus type 1.";  
RL J. Gen. Virol. 69:1531-1574 (1988).  
RN [2]  
RP REVISIONS.  
RX MEDLINE=92341080; PubMed=1321882;  
RA Dolan A., McKie E., McLean A.R., McGeoch D.J.;  
RT "Status of the ICP34.5 gene in herpes simplex virus type 1 strain  
RT 17.";  
RL J. Gen. Virol. 73:971-973 (1992).  
CC -1- FUNCTION: NOT KNOWN; HAS A PROFOUND EFFECT ON THE GROWTH OF THE  
CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.  
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X14112; -; NOT\_ANNOTATED\_CDS.  
CC PIR; JQ1682; JQ1682.  
CC Repeat.  
CC KW DOMAIN  
CC FT SEQUENCE 161 175 5 X 3 AA TANDEM REPEATS OF A-T-P.  
SQ SEQUENCE 248 AA; 26185 MW; AD334AF0CC9CABA65 CRC64;  
QY 437 DDDDDDD 443  
Db 75 DDDDDDD 81  
Query Match 1.5%; Score 7; DB 1; Length 248;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
RESULT 50  
YB95\_ARATH  
ID YB95\_ARATH STANDARD; PRT; 250 AA.  
AC O6W487;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein At2g39795, mitochondrial precursor.  
GN At2G39795 OR T517.20.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_Taxid=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=20083487; PubMed=10617197;  
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
RA Buell C.R., Ketchum K.A., Lee J.J., Rongning C.M., Koo H.L.,  
RA Moffat K.S., Cronin L.A., Shen M., Pat G., Van Aken S., Umayam L.,  
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,  
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,  
RA Nierman W.C., White O., Eissen J.A., Salzberg S.L., Frazer C.M.,  
RA Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
RT thaliana.";  
RL Nature 402:761-768 (1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;  
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the  
RT SGP consortium (Salk/Stanford/PGEC).";

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix (Potential).  
CC -1- SIMILARITY: BELONGS TO THE MAM33 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AC003000; AAM14855.1; -;  
CC DR EMBL; AY062757; AAL32835.1; -;  
CC DR InterPro; IPR003428; MAM33.  
CC DR Pfam; PF02330; MAM33; 1.  
CC KW Hypothetical protein; Mitochondrion; Transit peptide.  
CC FT TRANSIT 1 250 MITOCHONDRION (POTENTIAL).  
CC CHAIN ? 250 HYPOTHETICAL PROTEIN AT2G39795.  
SQ SEQUENCE 250 AA; 28061 MW; EB8884B390FE7D47 CRC64;  
QY 437 DDDDDDD 443  
Db 132 DDDDDDD 138  
Query Match 1.5%; Score 7; DB 1; Length 250;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
RESULT 51  
ICP3\_HSV1D  
ID ICP3\_HSV1D STANDARD; PRT; 252 AA.  
AC P37318;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE Infected cell protein ICP34.5 (Neurovirulence factor ICP34.5).  
GN ICP34.5.  
OS Herpes simplex virus (type 1 / strain CVG-2).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_Taxid=37106;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90156494; PubMed=2154589;  
RA Chou J., Roizman B.;  
RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted  
RT repeats, is conserved in several limited-passage isolates but not in  
RT strain 17syn+.";  
RL J. Virol. 64:1014-1020 (1990).  
CC -1- FUNCTION: NOT KNOWN; HAS A PROFOUND EFFECT ON THE GROWTH OF THE  
CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.  
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M33701; AAA5792.1; -;  
CC DR EMBL; M33701; AAA5792.1; -;  
CC KW Repeat; Late protein.  
CC FT DOMAIN 162 179 6 X 3 AA TANDEM REPEATS OF A-T-P.  
SQ SEQUENCE 252 AA; 26601 MW; A6E6D049FEBA79A CRC64;  
QY 437 DDDDDDD 443  
Db 132 DDDDDDD 138  
Query Match 1.5%; Score 7; DB 1; Length 252;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443  
 DB 76 DDDDDDD 82

RESULT 52  
 ASP\_PLAFS STANDARD; PRT; 253 AA.

DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Aspartic acid-rich protein precursor.  
 OS Plasmodium falciparum (isolate fcm17 / Senegal).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 NCBI\_TaxID=5845;

SEQUENCE FROM N.A.  
 MEDLINE=87270765; PubMed=3038111;  
 Lenstra R., D'Auriol L., Andrieu B., le Bras J., Galibert F.;  
 "Cloning and sequencing of Plasmodium falciparum DNA fragments  
 containing repetitive regions potentially coding for histidine-rich  
 proteins: identification of two overlapping reading frames."  
 Biochem. Biophys. Res. Commun. 146:368-377(1987).  
 -1- MISCELLANEOUS: THIS PROTEIN IS CODED ON THE REVERSE STRAND OF AN  
 HISTIDINE-RICH PROTEIN.  
 -1- SIMILARITY: SOME SIMILARITY WITH HUMAN SET PROTEIN.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M17028; AAA29620.1; -.  
 PIR; B29653; B29653.  
 InterPro; IPR002164; NAP\_family.  
 Pfam; PF00956; NAP; 1.  
 KMW Malaria; Antigen; Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 253 ASPARTIC ACID-RICH PROTEIN.  
 FT DOMAIN 204 253 ASP/GLU-RICH (ACIDIC).  
 SQ SEQUENCE 253 AA; 30248 MW; 83E585DE074B8504 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443  
 DB 218 DDDDDDD 224

RESULT 53  
 LI39 CAEEL STANDARD; PRT; 253 AA.

AC P34684;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Homeobox protein lin-39.  
 GN LIN-39 OR CEH-15 OR C07H6.7.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 NCBI\_TaxID=6239;

SEQUENCE FROM N.A.  
 STRAIN=Bristol N2;  
 MEDLINE=93327430; PubMed=8101475;  
 Clark S.G.; Chisholm A.D.; Horvitz H.R.;

RT "Control of cell fates in the central body region of C. elegans by  
 the homeobox gene lin-39."  
 RL Cell 74:43-55(1993).

SEQUENCE FROM N.A.  
 MEDLINE=93327429; PubMed=8101474;  
 Wang B.B., Mueller-Immergluck M.M., Austin J., Robinson N.T.,  
 Chisholm A.D., Kenyon C.;  
 "A homeotic gene cluster patterns the anteroposterior body axis of C.  
 elegans."  
 RL Cell 74:29-42(1993).

SEQUENCE FROM N.A.  
 STRAIN=Bristol N2;  
 Macri C., Vaudin M.;  
 Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 -1- FUNCTION: CONTROLS THE MIGRATION OF NEUROBLASTS AND THE  
 SUBSEQUENT MID-BODY REGION SPECIFIC DEVELOPMENT.  
 -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 -1- SIMILARITY: BELONGS TO THE ANTP HOMEBOX FAMILY.  
 "DEFORMED" SUBFAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; L19639; AAC37168.1; -.  
 EMBL; L19248; AAB04137.1; -.  
 EMBL; AC006605; AAK85445.1; -.  
 PIR; A40720; A40720.  
 HSP; P02833; 1SAN.  
 WormBep; C07H6.7; CE03975.  
 TRANSFAC; T03369; -.  
 InterPro; IPR001827; Antennapedia.  
 InterPro; IPR001356; Homeobox.  
 Pfam; PF00046; homeobox; 1.  
 PRINTS; PR00025; ANTENNAPEDIA.  
 PRINTS; PR00024; HOMEBOX.  
 PRODOM; PD000010; Homeobox; 1.  
 SMART; SM00389; HOX; 1.  
 PROSITE; PS00027; HOMEBOX\_1; 1.  
 PROSITE; PS50071; HOMEBOX\_2; 1.  
 PROSITE; PS00032; ANTENNAPEDIA; 1.  
 DR Homeobox; DNA-binding; Nuclear protein; Developmental protein.  
 KW Homeobox; DNA-binding; Nuclear protein; Developmental protein.  
 FT DOMAIN 20 36 POLY-SER.  
 FT DOMAIN 121 129 POLY-ASP.  
 FT DOMAIN 148 153 ANTP-TYPE HEXAPEPTIDE.  
 FT DNA\_BIND 165 224 HOMEBOX.  
 FT DOMAIN 228 248 PRO-RICH.  
 SQ SEQUENCE 253 AA; 27814 MW; 6B59DE9568B834A1 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 435 EGDGDDD 441  
 DB 119 EGDGDDD 125

RESULT 54

T2D7 RAT STANDARD; PRT; 253 AA.

AC Q62880;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Transcription initiation factor TFIIID 31 kDa subunit (TAFII-31)  
 (TAFII-32) (TAFII32) (Neuronal cell death related gene in neuron -7)

DE (DN-7).  
GN TAF9 OR TAF2G OR TAFI131.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97312553; PubMed=9168994;  
RA Aoki T., Koike T., Nakano T., Shibahara K., Nishimura H., Kikuchi H.,  
RA Honjo T.;  
RT "Rat TAFI131 gene is induced upon programmed cell death in  
RT differentiated PC12 cells deprived of NGF.";  
RL Blochem. Biophys. Res. Commun. 234:230-234(1997).  
CC -1- FUNCTION: TAFs ARE COMPONENTS OF THE TRANSCRIPTION FACTOR IID  
CC (TFIID) COMPLEX THAT ARE ESSENTIAL FOR MEDIATING REGULATION OF RNA  
CC POLYMERASE TRANSCRIPTION. TAFI131 IS A COACTIVATOR FOR THE P53  
CC PROTEIN. ALSO INTERACTS WITH THE ACIDIC TRANSACTIVATOR VIRAL  
CC PROTEIN 16 (VP16) AS WELL AS WITH THE GENERAL TRANSCRIPTION FACTOR  
CC TFIIB (BY SIMILARITY).  
CC -1- SUBUNIT: TFIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A  
CC NUMBER OF TBP-ASSOCIATED FACTORS (TAFs) (BY SIMILARITY). TAFI131  
CC AND P53 BIND TO EACH OTHER VIA AMINO ACID RESIDUES IN THE AMINO-  
CC TERMINAL DOMAIN OF P53 THAT ARE ESSENTIAL FOR TRANSCRIPTION (BY  
CC SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE TAF2G FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U40188; AAC53201.1; -.  
DR InterPro; IPR004822; Histone\_core.  
DR InterPro; IPR003162; TFIID-31.  
DR Pfam; PF02291; TFIID-31; 1.  
DR ProDom; PD01023; TFIID-31; 1.  
KW Transcription regulation; Nuclear protein.  
FT DOMAIN 238 251 POLY-ASP.  
SQ SEQUENCE 253 AA; 27620 MW; 42045091A9B94378 CRC64;  
Query Match 1.5%; Score 7; DB 1; Length 253;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
437 DDDDDDD 443  
|||||  
Db 238 DDDDDDD 244  
RESULT 55  
ICP3\_HSV1F  
ID ICP3\_HSV1F STANDARD; PRT; 263 AA.  
AC P08353;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE Infected cell protein ICP34.5 (Neurovirulence factor ICP34.5).  
GN ICP34.5.  
OS Herpes simplex virus (type 1 / strain F).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=10304;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86115412; PubMed=3003394;  
RA Chou J., Roizman B.;  
RT "The terminal a sequence of the herpes simplex virus genome contains  
RT the promoter of a gene located in the repeat sequences of the L

RT component.";  
RL J. Virol. 57:629-637(1986).  
RN [2]  
RP REVISIONS.  
RX MEDLINE=90156494; PubMed=2154589;  
RA Chou J., Roizman B.;  
RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted  
RT repeats, is conserved in several limited-passage isolates but not in  
RT strain 17syn+.";  
RL J. Virol. 64:1014-1020(1990).  
CC -1- FUNCTION: NOT KNOWN; HAS A PROFOUND EFFECT ON THE GROWTH OF THE  
CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.  
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M12240; AAA45794.1; ALT\_SEQ.  
DR EMBL; M33699; AAA45790.1; -.  
DR PIR; A27768; WMBE38.  
KW Repeat; Late protein.  
FT DOMAIN 161 190  
SQ SEQUENCE 263 AA; 27533 MW; F5084106A08A8CBI CRC64;  
OY 437 DDDDDDD 443  
|||||  
Db 75 DDDDDDD 81  
Query Match 1.5%; Score 7; DB 1; Length 263;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
437 DDDDDDD 443  
|||||  
Db 75 DDDDDDD 81  
RESULT 56  
OSTP\_CHICK  
ID OSTP\_CHICK STANDARD; PRT; 264 AA.  
AC P23498;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Osteopontin precursor (Bone sialoprotein 1).  
GN Spp1.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91236779; PubMed=2033080;  
RA Castagnola P., Bec P., Quarto R., Gennari M., Migliaccio G.,  
RA Cancedda R.;  
RT "cDNA cloning and gene expression of chicken osteopontin. Expression  
RT of osteopontin mRNA in chondrocytes is enhanced by trypsin treatment  
RT of cells.";  
RL J. Biol. Chem. 266:9944-9949(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91159433; PubMed=2001376;  
RA Moore M.A., Gotoh Y., Rafidi K., Gerstenfeld L.C.;  
RT "Characterization of a cDNA for chicken osteopontin: expression  
RT during bone development, osteoblast differentiation, and tissue  
RT distribution.";  
RL Biochemistry 30:2501-2508(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94192994; PubMed=8144023;



RA Rafidi K., Simikina I., Johnson E., Moore M.A., Gerstenfeld L.C.;  
 RT "Characterization of the chicken osteopontin-encoding gene";  
 RL Gene 140:163-169(1994).  
 CC -1- FUNCTION: BINDS TIGHTLY TO HYDROXYAPATITE. APPEARS TO FORM AN  
 INTEGRAL PART OF THE MINERALIZED MATRIX. PROBABLY IMPORTANT TO  
 CELL-MATRIX INTERACTION.  
 CC -1- SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3.  
 CC -1- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES.  
 CC -1- PTM: N- AND O-GLYCOSYLATED.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOPONTIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X56772; CAA40091.1; -  
 CC EMBL; M59182; AAA62729.1; -  
 CC EMBL; U01844; AAA18584.1; -  
 DR PIR; A37962; A37962.  
 DR PIR; A40019; A40019.  
 DR InterPro; IPR002038; Osteopontin.  
 DR Pfam; PF00865; Osteopontin; 1.  
 DR SMART; SM00017; OSTEO; 1.  
 DR PROSITE; PS00884; OSTEOPONTIN; 1.  
 KW Glycoprotein; Sialic acid; Bone; Cell adhesion; Phosphorylation;  
 KW Signal.  
 FT SIGNAL 1 16  
 FT CHAIN 17 264  
 FT SITE 132 134  
 FT CARBOHYD 106 106  
 FT CARBOHYD 109 109  
 FT CARBOHYD 204 204  
 FT CARBOHYD 242 242  
 FT CONFLICT 104 104  
 FT CONFLICT 216 216  
 FT CONFLICT 235 235  
 FT CONFLICT 240 240  
 SQ SEQUENCE 264 AA; 29162 MW; CA2F17DB70A92591 CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 437 DDDDDDD 443  
 |||||  
 99 DDDDDDD 105  
 RESULT 57  
 T2D7 HUMAN STANDARD; PRT; 264 AA.  
 ID T2D7 HUMAN  
 AC Q16594;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Transcription initiation factor TFIID 31 kDa subunit (TAFII-31)  
 DE (TAFII-32) (TAFII32).  
 DE TAF9 OR TAF2G OR TAFII31.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RX MEDLINE=95281615; PubMed=7761466;  
 RA Lu H., Levine A.J.;  
 RT "Human TAFII31 protein is a transcriptional coactivator of the p53  
 protein";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:5154-5158(1995).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95320160; PubMed=7597030;  
 RA Klemm R.D., Goodrich J.A., Zhou S., Tjian R.;  
 RT "Molecular cloning and expression of the 32-kDa subunit of human  
 TFIID reveals interactions with VP16 and TFIIB that mediate  
 transcriptional activation";  
 RT Proc. Natl. Acad. Sci. U.S.A. 92:5788-5792(1995).  
 RL [3]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Brain;  
 RX MEDLINE=95396764; PubMed=7667268;  
 RA Hisatake K., Ohta T., Takada R., Guernah M., Horikoshi M.,  
 Nakatani Y., Roeder R.G.;  
 RT "Evolutionary conservation of human  
 TATA-binding-protein-polypeptide-associated factors TAFII31 and TAFII80 and  
 interactions of TAFII80 with other TAFs and with general  
 transcription factors";  
 RT Proc. Natl. Acad. Sci. U.S.A. 92:8195-8199(1995).  
 CC -1- FUNCTION: TAFs ARE COMPONENTS OF THE TRANSCRIPTION FACTOR IID  
 (TFIID) COMPLEX THAT ARE ESSENTIAL FOR MEDIATING REGULATION OF RNA  
 POLYMERASE TRANSCRIPTION. TAFII31 IS A COACTIVATOR FOR THE P53  
 PROTEIN. ALSO INTERACTS WITH THE ACIDIC TRANSACTIVATOR VIRAL  
 PROTEIN 16 (VP16) AS WELL AS WITH THE GENERAL TRANSCRIPTION FACTOR  
 TFIIB.  
 CC -1- SUBUNIT: TFIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A  
 NUMBER OF TBP-ASSOCIATED FACTORS (TAFs). TAFII31 AND P53 BIND TO  
 EACH OTHER VIA AMINO ACID RESIDUES IN THE AMINO-TERMINAL DOMAIN OF  
 P53 THAT ARE ESSENTIAL FOR TRANSCRIPTION.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE TAF2G FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U25112; AAA91318.1; -  
 CC EMBL; U21858; AAC50153.1; -  
 CC EMBL; U30504; AAA84389.1; -  
 DR TRANSFAC; T02113; -  
 DR Genew; HGNC:11542; TAF9.  
 DR MIM; 600822; -  
 DR InterPro; IPR004822; Histone\_core.  
 DR InterPro; IPR003162; TFIID-31.  
 DR Pfam; PF02291; TFIID-31; 1.  
 DR ProDom; PD011023; TFIID-31; 1.  
 KW Transcription regulation; Nuclear protein.  
 FT DOMAIN 250 262  
 SQ SEQUENCE 264 AA; 28974 MW; 1925AEC65D6C84C7 CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 437 DDDDDDD 443  
 |||||  
 250 DDDDDDD 256  
 RESULT 58  
 EF1D XENLA STANDARD; PRT; 265 AA.  
 ID EF1D XENLA  
 AC P29693;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE Elongation factor 1-delta (EF-1-delta) (P36).  
 DE Xenopus laevis (African clawed frog).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipioidea; Pipidae;  
CC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Oocyte;  
RX MEDLINE=92375699; PubMed=1508694;  
RA Morales J., Cormier P., Mulner-Lorillon O., Poulhe R., Belle R.,  
RT "Molecular cloning of a new guanine nucleotide-exchange protein, EPI  
delta.";  
RL Nucleic Acids Res. 20:4091-4091 (1992).  
RN [2]  
RP PRELIMINARY SEQUENCE OF 15-27; 112-118; 186-192 AND 226-259.  
RX MEDLINE=91331996; PubMed=1869528;  
RA Janssen G.M.C., Morales J., Schipper A., Labbes J.C.,  
RA Mulner-Lorillon O., Belle R., Moeller W.,  
RT "A major substrate of maturation promoting factor identified as  
RT elongation factor 1 beta gamma delta in Xenopus laevis.";  
RL J. Biol. Chem. 266:14885-14888 (1991).  
CC -1- FUNCTION: EF-1-BETA AND EF-1-DELTA STIMULATE THE EXCHANGE OF  
CC GDP BOUND TO EF-1-ALPHA TO GTP.  
CC -1- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,  
CC DELTA, AND GAMMA.  
CC -1- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X66837; CAA47313.1; -.  
DR PIR; A39331; A39331.  
DR PIR; S24635; S24635.  
DR PIR; S26280; S26280.  
DR InterPro; IPR001326; EF1\_BD.  
DR Pfam; PF00736; EF1BD\_1.  
DR PROSITE; PS00824; EF1BD\_1; 1.  
DR PROSITE; PS00825; EF1BD\_2; 1.  
DR KW Elongation factor; Protein biosynthesis.  
SQ SEQUENCE 265 AA; 29237 MW; 61A1898EC3F9E402 CRC64;  
Query Match 1.5%; Score 7; DB 1; Length 265;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
437 DDDDDDD 443  
| | | | |  
Db 134 DDDDDDD 140  
RESULT 59  
HAX1\_MOUSE  
ID HAX1\_MOUSE STANDARD; PRT; 280 AA.  
AC O35387;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE HSI-associating protein X-1 (HAX-1) (HS1-binding protein).  
GN HAX1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Watanabe T., Takeshita H.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: DIRECTLY ASSOCIATES WITH HSI, THROUGH BINDING TO ITS N-  
CC TERMINAL REGION. MAY FUNCTION IN PROMOTING CELL SURVIVAL. MAY ALSO  
CC ASSOCIATE WITH CORFACIN/EMSI IN NONLYMPHOID CELLS (BY

CC SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: MAINLY IN MITOCHONDRION, BUT ALSO TO A  
CC LESSER EXTENT IN ENDOPLASMIC RETICULUM AND NUCLEAR ENVELOPE (BY  
CC SIMILARITY).  
CC -1- TISSUE SPECIFICITY: UBQUITOUS.  
CC -1- SIMILARITY: WEAK, TO THE NIP3 AND BCL-2 FAMILY PROTEINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF023482; AAB81081.1; -.  
DR MGD; MGI:1346319; Hax1.  
DR Mitochondrion.  
FT DOMAIN 30 44 ASP/GLU-RICH (HIGHLY ACIDIC).  
FT DOMAIN 115 280 INVOLVED IN HSI BINDING (BY SIMILARITY).  
SQ SEQUENCE 280 AA; 31654 MW; 3123979BA10B45D1 CRC64;  
Query Match 1.5%; Score 7; DB 1; Length 280;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDD 443  
| | | | |  
Db 30 DDDDDDD 36  
RESULT 60  
YNV9\_CAEEL  
ID YNV9\_CAEEL STANDARD; PRT; 289 AA.  
AC P34572; Q23407;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein T16H12.9 in chromosome III.  
GN T16H12.9.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Smith A., Berke M.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP REVISIONS.  
RA Durbin R.;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z30662; CAA83142.2; -.  
DR EMBL; Z47357; CAA83142.2; JOINED.  
DR EMBL; Z47357; CAA87428.2; -.  
DR EMBL; Z30662; CAA87428.2; JOINED.  
DR PIR; S42381; S42381.  
DR WormPep; T16H12.9; CE28843.  
DR Hypothetical protein.  
KW DOMAIN 134 145 ASP-RICH.  
FT SEQUENCE 289 AA; 32186 MW; 425B5C6107108D7F CRC64;  
Query Match 1.5%; Score 7; DB 1; Length 289;

Best Local Similarity 100.0%; Pred. No. 41;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 134 DDQDDDD 140

RESULT 61

UBC3\_YEAST

ID UBC3\_YEAST STANDARD; PRT; 295 AA.

AC P14682;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Ubiquitin-conjugating enzyme E2-34 kDa (EC 6.3.2.19)

DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein) (Cell division

DE control protein 34).

GN UBC3 OR CDC34 OR DNA6 OR YDR054C OR YD9609.08C OR D4211.

OS Saccharomyces cerevisiae (Baker's yeast).

OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OS Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI\_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88321694; PubMed=2842867;

RA Goehl M.G., Yochem J., Jentsch S., McGrath J.P., Varshavsky A.,

RA Byers B.;

RT "The yeast cell cycle gene CDC34 encodes a ubiquitin-conjugating

RT enzyme."

RL Science 241:1331-1335(1988).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=96381250; PubMed=8789263;

RA Brandt P., Ramlow S., Otto B., Bloeker H.;

RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm

RT of Saccharomyces cerevisiae chromosome IV."

RL Yeast 12:85-90(1996).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;

RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;

RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO

CC OTHER PROTEINS. CAPABLE, IN VITRO, TO UBIQUITINATE HISTONE H2A.

CC -1- FUNCTION: MEDIATES THE INITIATION OF DNA REPLICATION (TRANSITION

CC OF G1 TO S PHASE IN CELL CYCLE).

CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +

CC diphosphate + protein N-ubiquityllysine.

CC -1- PATHWAY: Ubiquitin conjugation; second step.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- DOMAIN: THE ACIDIC C-TERMINAL EXTENSION IS ESSENTIAL FOR THE CELL

CC CYCLE FUNCTION.

CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR

CC UBIQUITIN-THIOLESTER FORMATION.

CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL; M21877; AAA5188.1; -

CC EMBL; X84162; CAA58970.1; -

CC EMBL; Z74350; CAA98872.1; -

CC EMBL; Z49209; CAA89083.1; -

CC PIR; A41241; A41241.

CC HSSP; O02159; 2UCZ.

CC SGD; S0002461; CDC34.

CC InterPro; IPR000608; UBQ\_conjugat.

DR Pfam; PF00179; UQ\_con; 1.  
DR ProDom; PD000461; UBQ\_conjugat; 1.  
DR SMART; SM00212; UBCC; 1.  
DR PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
DR PROSITE; PS50127; UBIQUITIN\_CONJUGAT\_2; 1.  
KW Ubl conjugation pathway; ligase; DNA replication; Nuclear protein;  
KW Cell cycle; Cell division; Multigene family.  
FT BINDING 95 95 UBIQUITIN (BY SIMILARITY).  
FT DOMAIN 191 289 ASP/GLU-RICH (ACIDIC).  
SQ SEQUENCE 295 AA; 34064 MW; 1CE3E0C3AB1436DC CRC64;

Query Match 1.5%; Score 7; DB 1; Length 295;

Best Local Similarity 100.0%; Pred. No. 42;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 255 DDDDDDD 261

RESULT 62

PEX5\_CANAL

ID PEX5\_CANAL STANDARD; PRT; 296 AA.

AC 074711;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Peroxisomal targeting signal receptor (Peroxisomal protein PAS10)

DE (Peroxin-5) (PTS1 receptor) (Fragment).

GN PEX5.

OS Candida albicans (Yeast).

OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OS Saccharomycetales; mitosporic Saccharomycetales; Candida.

OX NCBI\_TaxID=5476;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 64385 / 1001;

RA Navarro-Garcia F., Perez-Diaz R., Negro A., Pla J., Nombela C.;

RT "Cloning and sequence of a 3.835 kbp DNA fragment containing the HIS4

RT gene and a fragment of a PEX5-like gene from Candida albicans."

RL Yeast 14:1147-1157(1998).

CC -1- FUNCTION: BINDS TO THE C-TERMINAL PTS1-TYPE TRIPEPTIDE PEROXISOMAL

CC TARGETING SIGNAL (SKL-TYPE) AND PLAYS AN ESSENTIAL ROLE IN

CC PEROXISOMAL PROTEIN IMPORT (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: ITS DISTRIBUTION APPEARS TO BE DYNAMIC. IT

CC IS PROBABLY A CYCLING RECEPTOR FOUND MAINLY IN THE CYTOPLASM AND

CC AS WELL ASSOCIATED TO THE PEROXISOMAL MEMBRANE THROUGH A DOCKING

CC FACTOR (PEX13) (BY SIMILARITY).

CC -1- SIMILARITY: CONTAINS 8 TPR REPEATS.

CC -1- SIMILARITY: STRONG, TO OTHER PEROXISOMAL TARGETING SIGNAL

CC RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL; AJ003115; CAA05870.1; -

CC InterPro; IPR001440; TPR.

CC Pfam; PF00515; TPR; 4.

CC SMART; SM00028; TPR; 4.

CC Peroxisome; Repeat; TPR repeat; Transport; Protein transport.

CC

CC NON TER 1 1

CC REPEAT 1 1

CC REPEAT 34 33 TPR 2.

CC REPEAT 67 67 TPR 3.

CC REPEAT 105 105 TPR 4.

CC REPEAT 143 143 TPR 5.

CC REPEAT 144 144 TPR 6.

CC REPEAT 177 177 TPR 7.

CC REPEAT 211 211

CC

CC

```
FT REPEAT      212      245      TPR 8.
SQ SEQUENCE    296 AA; 33220 MW; 1F78FAE76395A0B CRC64;
Query Match      1.5%; Score 7; DB 1; Length 296;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 NSNRSEE 101
   |||||
Db 190 NSNRSEE 196

RESULT 63
MURB_BACSU
ID MURB_BACSU STANDARD; PRT; 303 AA.
AC P18579; P16669; P37581;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DT UDP-N-acetylglucosaminylglucosamine reductase (EC 1.1.1.158) (UDP-N-
acetylmutamate dehydrogenase).
MURB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90078133; PubMed=2556375;
RA Beall B., Lutkenhaus J.;
RT "Nucleotide sequence and insertional inactivation of a Bacillus
RT subtilis gene that affects cell division, sporulation, and
RL temperature sensitivity."
RL J. Bacteriol. 171:6821-6834(1989).
RN [2]
RP SEQUENCE OF 1-23 FROM N.A.
RX MEDLINE=92380484; PubMed=1387377;
RA Miyao A., Yoshimura A., Sato T., Yamamoto T., Theeragool G.,
RA Kobayashi Y.;
RT "Sequence of the Bacillus subtilis homolog of the Escherichia coli
RT cell-division gene murG."
RL Gene 118:147-148(1992).
RN [3]
RP SEQUENCE OF 250-303 FROM N.A.
RX MEDLINE=90078134; PubMed=2556376;
RA Harry E.J., Wake R.G.;
RT "Cloning and expression of a Bacillus subtilis division initiation
RT gene for which a homolog has not been identified in another
RT organism."
CC J. Bacteriol. 171:6835-6839(1989).
RN [4]
RP FUNCTION.
RX MEDLINE=96060848; PubMed=7590298;
RA Rowland S.L., Errington J., Wake R.G.;
RT "The Bacillus subtilis cell-division 135-137 degrees region contains
RT an essential ofr with significant similarity to murB and a
RT dispensable sbp gene."
RL Gene 164:113-116(1995).
CC -1- FUNCTION: CELL WALL FORMATION.
CC -1- CATALYTIC ACTIVITY: UDP-N-acetylmutamate + NADP(+) = UDP-N-
CC acetyl-3-O-(1-carboxyvinyl)-D-glucosamine + NADPH.
CC -1- COFACTOR: FAD.
CC -1- PATHWAY: Peptidoglycan biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE MURB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; M31827; AAA83969.1; -.
DR EMBL; D10602; BAA01455.1; -.
DR EMBL; M31800; AAA22392.1; -.
DR EMBL; Z99111; CAB13396.1; -.
DR PIR; S26500; S26500.
DR PIR; A43727; A43727.
DR PIR; PC1128; PC1128.
DR Subtilist; BG10228; murB.
DR InterPro; IPR003170; murB.
DR InterPro; IPR001575; Oxid_FAD_bind.
DR Pfam; PF01565; FAD_binding_4_1.
DR Pfam; PF02873; MurB_C_1.
DR TIGRFAMs; TIGR00179; murB; 1.
KW Peptidoglycan synthesis; Cell wall; Cell division; Oxidoreductase;
KW NADP; Flavoprotein; FAD; Complete proteome.
SQ SEQUENCE 303 AA; 32808 MW; C33F5367C84E8A6E CRC64;

Query Match      1.5%; Score 7; DB 1; Length 303;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 QELKERE 312
   |||||
Db 6 QELKERE 12

RESULT 64
CDK3_HUMAN
ID CDK3_HUMAN STANDARD; PRT; 305 AA.
AC Q00526;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DT Cell division protein kinase 3 (EC 2.7.1.-).
GN CDK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Fetal brain;
RA MEDLINE=92347325; PubMed=1639063;
RA Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C.,
RA Harlow E., Tsai L.-H.;
RT "A family of human cdc2-related protein kinases."
RL EMBO J. 11:2909-2917(1992).
CC -1- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH A YET UNKNOWN TYPE OF CYCLIN. CAN PHOSPHORYLATE
CC HISTONE H1.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```





```

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Singer J.D., Manning B.E., Formosa T.;
RT "Control of single-copy DNA replication requires genes that act in
RL ubiquitin metabolism.";
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96381250; PubMed=8789263;
RA Brandt P., Ramlow S., Otto B., Bloeker H.;
RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm
RL of Saccharomyces cerevisiae chromosome IV.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=5288c / AB972;
RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;
RT Submitted (May-1995) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: ACTS IN UBIQUITIN METABOLISM AND IS NECESSARY FOR THE
CONTROL OF SINGLE-COPY DNA REPLICATION.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U19857; AAA66522.1; -
DR EMBL; X84162; CAA58984.1; -
DR EMBL; Z74364; CAA98886.1; -
DR EMBL; Z49209; CAA89097.1; -
DR EMBL; Z46796; CAA86790.1; -
DR SGD; S0002475; DOS2.
FT DOMAIN 255 259 POLY-GLU.
FT DOMAIN 302 308 POLY-ASP.
SQ SEQUENCE 310 AA; 35964 MW; B9B318093E984713 CRC64;

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 310;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
| | | | |
302 DDDDDDD 308

RESULT 67
MK03_MOUSE STANDARD; PRT; 314 AA.
AC Q63844; G61531;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase 3 (EC 2.7.1.-) (Extracellular signal-
DE regulated kinase 1) (ERK-1) (Insulin-stimulated MAP2 kinase) (MAP
DE kinase 1) (MAPK 1) (P44-ERK1) (ERT2) (P44-MAPK) (Microtubule-
DE associated protein-2 kinase) (MNK1) (Fragments).
GN MAPK3 OR PRKM3 OR ERK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-10 FROM N.A.
RC TISSUE=Pre-B cell;
RX MEDLINE=92020947; PubMed=1717989;
RA Crews C.M., Alessandrini A.A., Erikson R.L.;
RT "Mouse Erk-1 gene product is a serine/threonine protein kinase that
RT has the potential to phosphorylate tyrosine.";
```

```

RL Proc. Natl. Acad. Sci. U.S.A. 88:8845-8849(1991).
RN [2]
RP SEQUENCE OF 11-314 FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=91369479; PubMed=1716439;
RA de Miguel C., Kligman D., Patel J., Detera-Wadleigh S.D.;
RT "Molecular analysis of microtubule-associated protein-2 kinase cDNA
RL from mouse and rat brain.";
RN [3]
RP SEQUENCE OF 105-143 FROM N.A.
RC STRAIN=CBA; TISSUE=Bone marrow;
RX MEDLINE=93185941; PubMed=8444355;
RA Ershler M.A., Nagorskaya T.V., Visser J.W.M., Belyavsky A.V.;
RT "Novel CDC2-related protein kinases produced in murine hematopoietic
RT stem cells.";
RL Gene 124:305-306(1993).
[4]
RP SEQUENCE OF 105-139 FROM N.A.
RX MEDLINE=93092802; PubMed=1459009;
RA Ershler M.A., Nagorskaya T.V., Visser J.W.M., Belyavsky A.V.;
RT "Identification of new protein kinase genes, similar to kinases of
RT the cdc2 family and expressed in murine hematopoietic stem cells.";
RL Dokl. Akad. Nauk SSSR 324:893-897(1992).
-1- FUNCTION: PHOSPHORYLATES MICROTUBULE-ASSOCIATED PROTEIN-2 (MAP2).
CC MYELIN BASIC PROTEIN (MBP), AND ELK-1; MAY PROMOTE ENTRY IN THE
CC CELL CYCLE.
CC -1- ENZYME REGULATION: ACTIVATED BY TYROSINE AND THREONINE
CC PHOSPHORYLATION.
CC -1- PTM: AUTOPHOSPHORYLATED ON THREONINE AND TYROSINE RESIDUES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; S58470; AAB19973.1; -
DR EMBL; X64605; CAA45889.1; -
DR HSSP; P27703; 2BRK.
DR MGD; MGI:1346859; Mapk3.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; PARTIAL.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
KW Phosphorylation.
FT NON_TER 1 1
FT DOMAIN 1 4 POLY-GLY.
FT NON_CONS 10 11
FT MOD_RES 137 137 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT MOD_RES 139 139 (PROBABLE).
FT MOD_RES 139 139 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT MOD_RES 139 139 (PROBABLE).
FT CONFLICT 112 112 T -> P (IN REF. 3 AND 4).
SQ SEQUENCE 314 AA; 36198 MW; F06085E047BF7C22 CRC64;

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 314;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 LKICDFG 153
| | | | |
Db 115 LKICDFG 121
```

RESULT 68  
ID KRAF\_MSV36 STANDARD; PRT; 323 AA.  
AC P00532;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Serine/threonine-protein kinase transforming protein raf (EC 2.7.1.-).  
GN v-RAF.  
OS Murine sarcoma virus 3611.  
OC Viruses; Retrovirdae; Retroviridae; Mammalian type C retroviruses.  
OX NCBI\_TaxID=11812;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84121298; PubMed=6320371;  
RA Kan N.C., Flordellis C.S., Mark G.E., Duesberg P.H., Papas T.S.;  
RT "A common onc gene sequence transduced by avian carcinoma virus MH2  
and by murine sarcoma virus 3611.";  
RL Science 223:813-816(1984).  
[2]  
RT SEQUENCE FROM N.A.  
RX MEDLINE=84172180; PubMed=6324342;  
RA Mark G.E., Rapp U.R.;  
RT "Primary structure of v-raf: relatedness to the src family of  
oncogenes.";  
RL Science 224:285-289(1984).  
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-RAF  
POLYPEPTIDE.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC MIL/RAF SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; K01691; AAA46579.1; ALT\_INIT.  
DR PIR; A00638; TVMVF6.  
DR HSSP; P12931; IFMK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00221; STYKC; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Polypeptide; Serine/threonine-protein kinase; Transferase; Oncogene;  
KW ATP-binding.  
FT DOMAIN 24 284 PROTEIN KINASE.  
FT NP BIND 30 38 ATP (BY SIMILARITY).  
FT BINDING 50 50 ATP (BY SIMILARITY).  
FT ACT SITE 143 143 BY SIMILARITY.  
SQ SEQUENCE 323 AA; 36883 MW; 52A5423A6E362F3 CRC64;  
  
Query Match 1.5%; Score 7; DB 1; Length 323;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 172 WMAPEVI 178  
| | | | |  
Db 186 WMAPEVI 192  
  
RESULT 69  
LUKE\_STAAU STANDARD; PRT; 323 AA.  
ID LUKE\_STAAU  
AC P31715;

DT 01-JUL-1993 (Rel. 26, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Leukocidin F subunit precursor (Gamma-hemolysin, H-gamma-I subunit).  
GN LUKE.  
OS Staphylococcus aureus.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=1280;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-66.  
RC STRAIN=MRSA NO. 4;  
RX MEDLINE=92246945; PubMed=1575738;  
RA Rahman A., Nariya H., Izaki K., Kato I., Kamio Y.;  
RT "Molecular cloning and nucleotide sequence of leukocidin F-component  
gene (lukF) from methicillin resistant Staphylococcus aureus.";  
RL Biochem. Biophys. Res. Commun. 184:640-646(1992).  
[2]  
RN SEQUENCE OF 27-85 AND 322-323.  
RP STRAIN=RIMD 310925;  
RC MEDLINE=93223836; PubMed=8467905;  
RX Kamio Y., Rahman A., Nariya H., Ozawa T., Izaki K.;  
RT "The two Staphylococcal bi-component toxins, leukocidin and gamma-  
hemolysin, share one component in common.";  
RL FEBS Lett. 321:15-18(1993).  
CC -1- FUNCTION: LEUKOCIDIN CAUSES CYTOTOXIC CHANGES IN POLYMONONUCLEAR  
LEUKOCYTES. GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.  
CC -1- SUBUNIT: LEUKOCIDIN CONSISTS OF TWO PROTEIN COMPONENTS: F AND S;  
GAMMA-HEMOLYSIN CONSISTS OF TWO PROTEIN COMPONENTS (H-GAMMA-I= F  
AND H-GAMMA-II).  
CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; S65052; AAC60446.1; -.  
DR PIR; JQ1530; JQ1530.  
DR InterPro; IPR001776; Aerolysin.  
DR InterPro; IPR001340; Hemolysn\_pore.  
DR InterPro; IPR003963; Staph\_bicn\_tkn.  
DR Pfam; PF01117; Aerolysin; 1.  
DR PRINTS; PR01468; BICOMENTOXIN.  
DR TIGRFAMs; TIGR01002; hlyII; 1.  
KW Hemolysis; Toxin; Signal.  
FT SIGNAL 1 25  
FT CHAIN 26 323 LEUKOCIDIN F SUBUNIT.  
SQ SEQUENCE 323 AA; 36567 MW; A5951CF2AAB6C7E8 CRC64;  
  
Query Match 1.5%; Score 7; DB 1; Length 323;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 57 LSVLSHR 63  
| | | | |  
Db 250 LSVLSHR 256  
  
RESULT 70  
HLGB\_STAAU STANDARD; PRT; 325 AA.  
ID HLGB\_STAAU  
AC Q07226;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Gamma-hemolysin component B precursor (H-gamma-1) (H-gamma-1).  
GN HLGB OR SA2209 OR MW2344.  
OS Staphylococcus aureus (strain N315).  
OS Staphylococcus aureus (strain MW2), and  
OS Staphylococcus aureus.



OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=158879, 196620, 1280;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=N315;  
RX MEDLINE=21311952; PubMed=11418146;  
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
Mizutani-U. Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,  
"Whole genome sequencing of methicillin-resistant Staphylococcus  
aureus";  
RL Lancet 357:1225-1240(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MM2;  
RX MEDLINE=22040717; PubMed=12044378;  
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,  
Yamamoto K., Hiramatsu K.,  
"Genome and virulence determinants of high virulence community-  
acquired MRSA";  
RL Lancet 359:1819-1827(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93138806; PubMed=8423103;  
RA Cooney J.C., Klenle Z., Foster T.J., O'Toole P.W.,  
"The gamma-hemolysin locus of Staphylococcus aureus comprises three  
linked genes, two of which are identical to the genes for the F and S  
components of leukocidin";  
RL Infect. Immun. 61:768-771(1993).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89310396; PubMed=3075655;  
RA Cooney J.C., Mulvey M., Arbuthnot J., Foster T.,  
"Molecular cloning and genetic analysis of the determinant for gamma-  
lysin, a two-component toxin of Staphylococcus aureus";  
RL J. Gen. Microbiol. 134:2179-2188(1988).  
CC -1- FUNCTION: GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.  
CC -1- SUBUNIT: CONSISTS OF TWO PROTEIN COMPONENTS (H-GAMMA-I AND -II).  
CC H-GAMMA-I IS ALSO THE F SUBUNIT OF LEUKOCIDIN.  
CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AP003137; BAB43511.1; -  
DR EMBL; AP004830; BAB96209.1; -  
DR EMBL; L01055; AAA26639.1; -  
DR HSSP; P09616; 7AHL.  
DR InterPro; IPR001776; Aerolysin.  
DR InterPro; IPR001340; Hemlysin pore.  
DR InterPro; IPR003963; Staph\_bicn\_txn.  
DR Pfam; PF01117; Aerolysin; 1.  
DR PRINTS; PRO1468; BICOMPNTOXIN.  
DR TIGRfam; TIGR01002; hlyII; 1.  
KW Hemolysis; Toxin; Signal; Complete proteome.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 325 GAMMA-HEMOLYSIN COMPONENT B.  
SQ SEQUENCE 325 AA; 36711 MM; 082999EB9339B6A0 CRC64;

OY 57 LSVLSHR 63  
DB 252 LSVLSHR 258

RESULT 71  
KRAB\_MOUSE  
ID KRAB\_MOUSE STANDARD; PRT; 328 AA.  
AC P28028;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE B-Raf proto-oncogene serine/threonine-protein kinase (EC 2.7.1.-)  
DE (Fragment).  
GN BRAF OR B-Raf.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91271351; PubMed=2052597;  
RA Miki T., Fleming T.P., Crescenzi M., Molloy C.J., Blam S.B.,  
Reynolds S.H., Aaronson S.A.,  
"Development of a highly efficient expression cDNA cloning system:  
application to oncogene isolation";  
Proc. Natl. Acad. Sci. U.S.A. 88:5167-5171(1991).  
RT -1- FUNCTION: INVOLVED IN THE TRANSDUCTION OF MITOGENIC SIGNALS  
FROM THE CELL MEMBRANE TO THE NUCLEUS.  
CC -1- FUNCTION: MAY PLAY A ROLE IN THE POSTSYNAPTIC RESPONSES OF  
HIPPOCAMPAL NEURON.  
CC -1- DISEASE: PARTICIPATES IN A CHROMOSOMAL TRANSLOCATION THAT PRODUCES  
A T(11;22)(p11;p23) REARRANGEMENT.  
CC A T(11;22)(p11;p23) REARRANGEMENT ISOLATED FROM A FURFURAL-  
INDUCED HEPATOMA.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC MIL/RAF SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M64429; AAA37320.1; ALT\_INIT.  
DR PIR; A40951; TVMSBF.  
DR HSSP; P08631; 1AD5.  
DR MGD; MGI:88190; Braf.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00221; STYK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; Proto-oncogene;  
FT ATP-binding; Chromosomal translocation.  
FT NON\_TER 1 1  
FT DOMAIN 19 279 PROTEIN KINASE.  
FT NP\_BIND 25 33 ATP (BY SIMILARITY).  
FT BINDING 45 45 ATP (BY SIMILARITY).  
FT ACT\_SITE 138 138 BY SIMILARITY.  
SQ SEQUENCE 328 AA; 36986 MM; 67A2EBFB78A78E3D CRC64;

Query Match 1.5%; Score 7; DB 1; Length 328;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 130 IHRDLKS 136  
|||||



Db 135 IHRDLKS 141

RESULT 72

AAIP WHEAT STANDARD; PRT; 332 AA.

AC Q02066;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Abciscic acid-inducible protein kinase (EC 2.7.1.-) (Fragment).

OS Triticum aestivum (Wheat).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;

OC Triticaceae; Triticum.

OX NCBI\_TaxID=4565;

CC [1]

CC SEQUENCE FROM N.A.

CC STRAIN=cv. Brevor; TISSUE=Seed;

CC MEDLINE=93066204; PubMed=1438207;

CC "Isolation of a wheat cDNA clone for an absciscic acid-inducible

CC transcript with homology to protein kinases."

CC Proc. Natl. Acad. Sci. U.S.A. 89:10183-10187(1992).

CC -1- FUNCTION: IS INVOLVED IN WATER-STRESS RESPONSES.

CC -1- INDUCTION: BY ABSICISIC ACID AND DEHYDRATION.

CC -1- PTM: AUTOPHOSPHORYLATED (POTENTIAL).

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; M94726; AAA96325.1; -

CC PIR; A46408; A46408.

CC HSSP; Q63450; 1A06.

CC InterPro; IPR000719; Euk\_pkinase.

CC InterPro; IPR002290; Ser\_thr\_pkinase.

CC Pfam; PF00069; pkinase; 1.

CC ProDom; PD000001; Euk\_pkinase; 1.

CC SMART; SM00220; S\_TKc; 1.

CC DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; PARTIAL.

CC DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.

CC DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.

CC DR Transferase; Serine/threonine-protein kinase; ATP-binding;

CC Phosphorylation.

CC NON\_TER 1 1

CC DOMAIN 1 250 PROTEIN KINASE.

CC FT NP\_BIND <1 8 ATP (BY SIMILARITY).

CC FT BINDING 23 23 ATP (BY SIMILARITY).

CC FT ACT\_SITE 113 113 BY SIMILARITY.

CC SQ SEQUENCE 332 AA; 37516 MW; E19ACAE1C0F0CE37 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 332;

Best Local Similarity 100.0%; Pred. No. 47;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 LKICDFG 153

Db 129 LKICDFG 135

RESULT 73

CDC2\_CAEEL STANDARD; PRT; 332 AA.

ID CDC2\_CAEEL

AC P34556;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34 protein

DE kinase).

GN NCC-1 OR T05G5.3.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

CC [1]

CC SEQUENCE FROM N.A.

CC STRAIN=Bristol N2;

CC MEDLINE=94150718; PubMed=7906398;

CC Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,

CC Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,

CC Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A., Jier M.,

CC Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,

CC Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,

CC Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,

CC Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,

CC Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,

CC Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,

CC Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,

CC Wohldman P.,

CC "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

CC elegans."

CC Nature 368:32-38(1994).

CC [2]

CC SEQUENCE FROM N.A.

CC STRAIN=Bristol N2;

CC Ferraz C., Thierry-Mieg D., le Peuch C.J.,

CC Submitted (SEP-1992) to the EMBL/GenBank/DBJ databases.

CC [3]

CC SEQUENCE FROM N.A.

CC MEDLINE=95131956; PubMed=7830726;

CC Mori H., Palmer R.E., Sternberg P.W.,

CC "The identification of a Caenorhabditis elegans homolog of p34cdc2

CC kinase."

CC Mol. Gen. Genet. 245:781-786(1994).

CC [4]

CC SEQUENCE FROM N.A.

CC STRAIN=Bristol N2;

CC MEDLINE=99225469; PubMed=10207147;

CC Boxem M., Srinivasan D.G., van den Heuvel S.,

CC "The Caenorhabditis elegans gene ncc-1 encodes a cdc2-related kinase

CC required for M phase in meiotic and mitotic cell divisions, but not

CC for S phase."

CC Development 126:2227-2239(1999).

CC -1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC

CC CELL CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE

CC AND MITOSIS. P34 IS A COMPONENT OF THE KINASE COMPLEX THAT

CC PHOSPHORYLATES THE REPETITIVE CARBOXYL-TERMINUS OF RNA

CC POLYMERASE II (BY SIMILARITY).

CC -1- ENZYME REGULATION: PHOSPHORYLATION INACTIVATES THE ENZYME

CC (BY SIMILARITY).

CC -1- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH A REGULATORY

CC SUBUNIT AND WITH A CYCLIN.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC CDC2/CDKX SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; Z27079; CAA81590.1; -

CC DR EMBL; X68384; CAA48455.1; -

CC DR EMBL; S75262; AAC60520.1; -

CC DR EMBL; AF129109; AAD37119.1; -

CC PIR; S26572; S26572.

CC PIR; S41003; S41003.

CC HSSP; P24941; 1CKP.

DR WormPep; T05G5.3; CE00315.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
Cell cycle; Cell division; Mitosis; Phosphorylation.  
FT DOMAIN 22 312 PROTEIN\_KINASE.  
FT NP\_BIND 28 36 ATP (BY SIMILARITY).  
FT BINDING 51 51 ATP (BY SIMILARITY).  
FT ACT\_SITE 146 146 BY SIMILARITY.  
FT MOD\_RES 32 32 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 33 33 PHOSPHORYLATION (BY SIMILARITY).  
FT CONFLICT 176 176 R -> P (IN REF. 3).  
SQ SEQUENCE 332 AA; 38295 MW; B6297E92949C8206 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 332;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRLDK 135  
Db 142 VIHRLDK 148

## RESULT 74

Y713\_CHLMU STANDARD; PRT; 332 AA.  
ID Y713\_CHLMU  
AC Q9PJM3;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein TC0713.  
GN TC0713.  
OS Chlamydia muridarum.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=83560;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=MOPn / N199;  
RX MEDLINE=20150255; Pubmed=10684935;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,  
Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
Winn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,  
Eisen J., Fraser C.M.;  
RT "Genome sequences of Chlamydia trachomatis MOPn and Chlamydia  
RT pneumoniae AR39.";  
RL Nucleic Acids Res. 28:1397-1406(2000).  
CC -1- SIMILARITY: BELONGS TO THE UPF0158 FAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL; AE002339; AAF39526.1; -

DR TIGR; TC0713; -

DR InterPro; IPR005361; UPF0158.

DR Pfam; PF03682; UPF0158; 1.  
KW Hypothetical protein; Complete proteome.

FT DOMAIN 301 316 ASP/GLU-RICH (ACIDIC).  
FT DOMAIN 308 316 POLY-ASP.

SO SEQUENCE 332 AA; 39202 MW; 27527DC5E6ABDEA4 CRC64;  
Query Match 1.5%; Score 7; DB 1; Length 332;

Best Local Similarity 100.0%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDD 443  
Db 308 DDDDDDD 314

## RESULT 75

SPT2\_YEAST  
ID SPT2\_YEAST STANDARD; PRT; 333 AA.  
AC P06843;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE SPT2 protein (Negative regulator of Ty transcription).  
GN SPT2 OR SPM2 OR SIN1 OR YER161C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=85267670; Pubmed=2991744;  
RA Roeder G.S., Beard C., Smith M., Keranen S.;  
RT "Isolation and characterization of the SPT2 gene, a negative  
RT regulator of Ty-controlled yeast gene expression.";  
RL Mol. Cell. Biol. 5:1543-1553(1985).  
RN [2]

RP SEQUENCE FROM N.A.  
RX MEDLINE=91304406; Pubmed=2072912;  
RA Kruger W., Herzkowitz I.;  
RT "A negative regulator of HO transcription, SIN1 (SPT2), is a  
RT nonspecific DNA-binding protein related to HMG1.";  
RL Mol. Cell. Biol. 11:4135-4146(1991).  
RN [3]

RP SEQUENCE FROM N.A.  
RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,  
RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;  
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.  
RN [4]

RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972;  
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,  
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,  
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,  
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CHROMATIN PROTEIN. GLOBAL REGULATORY PROTEIN THAT PLAYS  
CC POSITIVE AS WELL AS NEGATIVE REGULATORY ROLES IN TRANSCRIPTION.

CC -1- SUBUNIT: INTERACTS WITH SAPI AND CDC23.  
CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- SIMILARITY: PARTIAL (DNA-BINDING MOTIF) TO HMG PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL; M1165; AAA35083.1; -

DR EMBL; U18917; AAB64688.1; -

DR PIR; A23438; A23438.

DR PIR; S30813; S30813.

DR SCD; S0000963; SPT2.  
KW DNA-binding; Activator; Repressor; Transcription regulation;  
KW Nuclear protein.  
SO SEQUENCE 333 AA; 38551 MW; 60C3BED7EAA99535 CRC64;

Sun May 4 10:01:08 2003

us-09-757-982-5.oligo.rsp

Page 40

Query Match 1.5%; Score 7; DB 1; Length 333;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 446 EEDNDMD 452  
Db 232 EEDNDMD 238

Search completed: May 1, 2003, 20:53:44  
Job time : 44 secs

---

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:50:11 ; Search time 89 Seconds

(without alignments)  
1053.387 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455

Sequence: 1 MSSLGASFVQIKFDLQFFE.....GDDDDDDGEEEDNDMDNSE 455

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Hit size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

SPTREMBL\_21.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rv1rus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description        |
|------------|-------|-------------|--------|-----------|--------------------|
| 1          | 455   | 100.0       | 455    | 4 Q9HCC4  | Q9hcc4 homo sapien |
| 2          | 331   | 72.7        | 800    | 4 Q9HDD2  | Q9hdd2 homo sapien |
| 3          | 331   | 72.7        | 800    | 4 Q9HCC5  | Q9hcc5 homo sapien |
| 4          | 331   | 72.7        | 800    | 4 Q9NYL2  | Q9nyl2 homo sapien |
| 5          | 331   | 72.7        | 800    | 4 Q9NYE9  | Q9nye9 homo sapien |
| 6          | 166   | 36.5        | 454    | 11 Q9ESL3 | Q9esl3 mus musculu |
| 7          | 166   | 36.5        | 802    | 11 Q9ESL4 | Q9esl4 mus musculu |
| 8          | 166   | 36.5        | 371    | 13 Q9OZY8 | Q9ozy8 brachydanio |
| 9          | 11    | 10.8        | 371    | 13 Q9OZY8 | Q9ozy8 brachydanio |
| 10         | 11    | 2.4         | 855    | 5 Q01700  | Q01700 caenorhabdi |
| 11         | 11    | 2.4         | 977    | 5 Q9VW24  | Q9vw24 drosophila  |
| 12         | 9     | 2.0         | 141    | 10 Q9SMA2 | Q9sma2 oryza sativ |
| 13         | 9     | 2.0         | 205    | 5 Q8ST12  | Q8st12 dictyosteli |
| 14         | 9     | 2.0         | 233    | 4 O15318  | O15318 homo sapien |
| 15         | 9     | 2.0         | 297    | 12 Q98182 | Q98182 moluscum c  |
| 16         | 9     | 2.0         | 408    | 10 Q39585 | Q39585 chlamydomon |
|            |       |             | 453    | 10 Q855D1 | Q855d1 oryza sativ |

|    |     |      |           |                    |
|----|-----|------|-----------|--------------------|
| 17 | 2.0 | 460  | 10 Q855G9 | Q855g9 oryza sativ |
| 18 | 2.0 | 502  | 11 Q91YV1 | Q91yv1 mus musculu |
| 19 | 2.0 | 502  | 11 Q91YR0 | Q91yr0 mus musculu |
| 20 | 2.0 | 504  | 11 Q63559 | Q63559 rattus norv |
| 21 | 2.0 | 607  | 10 Q9FG45 | Q9fg45 arabidopsis |
| 22 | 2.0 | 611  | 10 Q9LKB4 | Q9lkb4 arabidopsis |
| 23 | 2.0 | 616  | 5 Q9GRH2  | Q9grh2 sycon rapha |
| 24 | 2.0 | 652  | 16 Q9A0J5 | Q9a0j5 streptococc |
| 25 | 2.0 | 764  | 3 Q12500  | Q12500 saccharomyc |
| 26 | 2.0 | 789  | 2 Q9XBP8  | Q9xbp8 myxococcus  |
| 27 | 2.0 | 856  | 10 Q9ZSM8 | Q9zsm8 arabidopsis |
| 28 | 2.0 | 898  | 10 Q04246 | Q04246 arabidopsis |
| 29 | 2.0 | 1059 | 3 Q9P8G2  | Q9p8g2 candida alb |
| 30 | 2.0 | 1472 | 5 Q9U5A8  | Q9u5a8 bombyx mori |
| 31 | 2.0 | 1978 | 5 Q8TSH2  | Q8tsh2 anopheles g |
| 32 | 2.0 | 2391 | 5 Q27732  | Q27732 plasmodium  |
| 33 | 1.8 | 101  | 4 Q15452  | Q15452 homo sapien |
| 34 | 1.8 | 106  | 12 Q9J872 | Q9j872 spodoptera  |
| 35 | 1.8 | 157  | 12 Q55500 | Q55500 human cytom |
| 36 | 1.8 | 161  | 12 Q9J867 | Q9j867 spodoptera  |
| 37 | 1.8 | 161  | 12 Q919Q2 | Q919q2 culex nigri |
| 38 | 1.8 | 162  | 16 Q9HWQ2 | Q9hwq2 pseudomonas |
| 39 | 1.8 | 180  | 3 Q9P840  | Q9p840 candida alb |
| 40 | 1.8 | 183  | 4 Q00193  | Q00193 homo sapien |
| 41 | 1.8 | 184  | 10 Q42341 | Q42341 arabidopsis |
| 42 | 1.8 | 194  | 12 Q9QAM0 | Q9qam0 white spot  |
| 43 | 1.8 | 194  | 12 Q91LE7 | Q91le7 white spot  |
| 44 | 1.8 | 194  | 12 Q8QTE0 | Q8qte0 white spot  |
| 45 | 1.8 | 198  | 10 Q9C7Y9 | Q9c7y9 arabidopsis |
| 46 | 1.8 | 204  | 10 Q39756 | Q39756 fagus sylv  |
| 47 | 1.8 | 206  | 10 Q9SWT7 | Q9swt7 hordeum vul |
| 48 | 1.8 | 209  | 10 Q945P0 | Q945p0 nicotiana t |
| 49 | 1.8 | 211  | 17 Q28960 | Q28960 archaeoglob |
| 50 | 1.8 | 213  | 11 Q91VH4 | Q91vh4 mus musculu |
| 51 | 1.8 | 218  | 10 Q8RZR9 | Q8rzt9 oryza sativ |
| 52 | 1.8 | 229  | 16 Q8X4P6 | Q8x4p6 escherichia |
| 53 | 1.8 | 237  | 2 Q50439  | Q50439 mycobacteri |
| 54 | 1.8 | 246  | 10 Q9AS90 | Q9as90 oryza sativ |
| 55 | 1.8 | 250  | 10 Q9M1H5 | Q9m1h5 arabidopsis |
| 56 | 1.8 | 258  | 5 Q9GRQ9  | Q9grq9 leishmania  |
| 57 | 1.8 | 259  | 11 Q9CXN1 | Q9cxl1 mus musculu |
| 58 | 1.8 | 260  | 10 Q9SIS0 | Q9s1s0 arabidopsis |
| 59 | 1.8 | 277  | 10 Q948F2 | Q948f2 oryza sativ |
| 60 | 1.8 | 279  | 11 Q64125 | Q64125 mus sp. ear |
| 61 | 1.8 | 292  | 5 Q02445  | Q02445 helobdella  |
| 62 | 1.8 | 293  | 4 Q9NMX4  | Q9nmx4 homo sapien |
| 63 | 1.8 | 297  | 5 Q17319  | Q17319 caenorhabdi |
| 64 | 1.8 | 302  | 10 Q64936 | Q64936 arabidopsis |
| 65 | 1.8 | 302  | 12 Q8VB23 | Q8vb23 white spot  |
| 66 | 1.8 | 307  | 5 Q9NKD3  | Q9nkd3 drosophila  |
| 67 | 1.8 | 317  | 11 P97827 | P97827 rattus norv |
| 68 | 1.8 | 318  | 13 Q9PTX7 | Q9ptx7 letheuteron |
| 69 | 1.8 | 318  | 17 Q27264 | Q27264 methanobact |
| 70 | 1.8 | 321  | 12 Q913E5 | Q913e5 white spot  |
| 71 | 1.8 | 321  | 12 Q8VAL2 | Q8val2 white spot  |
| 72 | 1.8 | 321  | 12 Q91L82 | Q91l82 white spot  |
| 73 | 1.8 | 325  | 10 Q8S667 | Q8s667 oryza sativ |
| 74 | 1.8 | 326  | 10 Q8S672 | Q8s672 oryza sativ |
| 75 | 1.8 | 335  | 5 Q8TIF9  | Q8tif9 dictyosteli |
| 76 | 1.8 | 351  | 4 Q9COP9  | Q9cof9 homo sapien |
| 77 | 1.8 | 354  | 4 Q9BXC3  | Q9bxc3 homo sapien |
| 78 | 1.8 | 359  | 10 Q8RXA9 | Q8rxax zea mays (m |
| 79 | 1.8 | 365  | 5 Q26776  | Q26776 trypanosoma |
| 80 | 1.8 | 368  | 10 Q8VZN7 | Q8vzn7 arabidopsis |
| 81 | 1.8 | 374  | 5 Q9SGW5  | Q9sgw5 arabidopsis |
| 82 | 1.8 | 379  | 10 Q76381 | Q76381 caenorhabdi |
| 83 | 1.8 | 379  | 10 Q49391 | Q49391 arabidopsis |
| 84 | 1.8 | 380  | 10 Q9LV94 | Q9lv94 arabidopsis |
| 85 | 1.8 | 395  | 11 Q9JLD4 | Q9jld4 rattus norv |
| 86 | 1.8 | 396  | 11 Q923V6 | Q923v6 rattus norv |
| 87 | 1.8 | 414  | 5 Q8T030  | Q8t030 drosophila  |
| 88 | 1.8 | 415  | 10 Q9MAV2 | Q9mav2 arabidopsis |
| 89 | 1.8 | 416  | 4 Q9P289  | Q9p289 homo sapien |



|     |   |     |     |    |        |                    |
|-----|---|-----|-----|----|--------|--------------------|
| 90  | 8 | 1.8 | 416 | 11 | Q99UT2 | Q99Jt2 mus musculu |
| 91  | 8 | 1.8 | 422 | 10 | Q40543 | Q40543 nicotiana t |
| 92  | 8 | 1.8 | 426 | 10 | Q9SZM8 | Q9SZM8 arabidopsis |
| 93  | 8 | 1.8 | 434 | 5  | Q23994 | Q23994 drosophila  |
| 94  | 8 | 1.8 | 434 | 13 | Q91B76 | Q91B76 pagrus majo |
| 95  | 8 | 1.8 | 436 | 10 | Q9LFV3 | Q9LFV3 arabidopsis |
| 96  | 8 | 1.8 | 438 | 10 | Q9M3H8 | Q9M3H8 cicer ariet |
| 97  | 8 | 1.8 | 439 | 5  | O16303 | O16303 caenorhabdi |
| 98  | 8 | 1.8 | 443 | 13 | Q91863 | Q91863 fugu rubrip |
| 99  | 8 | 1.8 | 445 | 11 | Q60877 | Q60877 mus musculu |
| 100 | 8 | 1.8 | 447 | 13 | Q8QGV6 | Q8QGV6 xenopus lae |

ALIGNMENTS

RESULT 1  
Q9HCC4 PRELIMINARY; PRT; 455 AA.  
ID Q9HCC4  
AC 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE MLTK-beta (Similar to sterile-alpha motif and leucine zipper  
DE containing kinase AZK) (Mixed lineage kinase) (Mixed lineage  
DE kinase-related kinase MRK-beta).  
GN MLTK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gotoh I., Adachi M., Nishida E.;  
RT "Identification and Characterization of a Novel MAP Kinase Kinase  
RT Kinase, MLTK.";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=COLON;  
RA Strausberg R.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Acton S.;  
RT "MLK-mixed lineage kinase.";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=21950776; Pubmed=11836244;  
RT Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;  
RT "MRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in  
RT gamma-Radiation-induced Cell Cycle Arrest.";  
RL J. Biol. Chem. 277:13873-13882(2002).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AB049734; BAB16445.1; -;  
DR EMBL; BC001401; AAH01401.1; -;  
DR EMBL; AF325454; AAK11615.1; -;  
DR EMBL; AF480462; AAL85892.1; -;  
DR HSSP; P12931; IFMK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00221; STYK; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00219; TYRK; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

| SEQ | SEQUENCE  | 455 AA;  | 51582 MM;    | E87DB84A4D58B752 CRC64; |
|-----|---|--|--------------|-------------------------|
| QY  | 1 MSLSGASFWQIKRDDLQFFENCGGGSGSVYRAKWSIQDKEVAVKLLIKEAEILSVL      | 100.0%;  | Score 455;   | DB 4; Length 455;       |
| Db  | 1 MSLSGASFWQIKRDDLQFFENCGGGSGSVYRAKWSIQDKEVAVKLLIKEAEILSVL      | Best Local Similarity 100.0%;                                | Pred. No. 0; |                         |
| QY  | 61 SHRNIIQFYGVILEPPNYGIVTEYASLGLDYINSNRSEEMDMHIMTWATDVAKGMHY    | Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |              |                         |
| Db  | 61 SHRNIIQFYGVILEPPNYGIVTEYASLGLDYINSNRSEEMDMHIMTWATDVAKGMHY    |  |              |                         |
| QY  | 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIQS |  |              |                         |
| Db  | 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIQS |  |              |                         |
| QY  | 181 LPVSECTDTSYSGVVLWEMLTREVPFKGLBGLQVAVLVKNERLTIPSSCPRSFABLLH  |  |              |                         |
| Db  | 181 LPVSECTDTSYSGVVLWEMLTREVPFKGLBGLQVAVLVKNERLTIPSSCPRSFABLLH  |  |              |                         |
| QY  | 241 QCWEADAKKRPSFKQIISLESMSNDTSLPDKNSFLHNKAEMRCETATLERLKLBERD   |  |              |                         |
| Db  | 241 QCWEADAKKRPSFKQIISLESMSNDTSLPDKNSFLHNKAEMRCETATLERLKLBERD   |  |              |                         |
| QY  | 301 LSFKEQELKERERRLKWEQKLTQSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI  |  |              |                         |
| Db  | 301 LSFKEQELKERERRLKWEQKLTQSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI  |  |              |                         |
| QY  | 361 TATSNGEHGMMNPSLOAMLMGFDFISMNKAGAVHSGMOINMOAKONSKTTSKRGRK    |  |              |                         |
| Db  | 361 TATSNGEHGMMNPSLOAMLMGFDFISMNKAGAVHSGMOINMOAKONSKTTSKRGRK    |  |              |                         |
| QY  | 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455                     |  |              |                         |
| Db  | 421 KVNMAIGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455                     |  |              |                         |

RESULT 2  
Q9HDD2 PRELIMINARY; PRT; 800 AA.  
ID Q9HDD2  
AC 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Placible mixed lineage kinase protein.  
GN MLTKLAK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LYMPHOID ORGAN;  
RA Abe Y., Ueda N.;  
RT "Placible Mixed-lineage kinase derived from IAK cell.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AB030034; BAB12040.1; -;  
DR HSSP; P12931; IFMK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR001660; SAM.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00454; SAM; 1.  
DR SMART; SM00221; STYK; 1.  
DR SMART; SM00220; S\_TKC; 1.

DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91155 MW; B2814509EC54B07A CRC64;

Query Match 72.7%; Score 331; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWISQDKEVAVKLLKIEKAEILSVL 60  
 DB 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWISQDKEVAVKLLKIEKAEILSVL 60  
 QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYINSNRSEEMDMDHMTWATDVAKGMHY 120  
 DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYINSNRSEEMDMDHMTWATDVAKGMHY 120  
 QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTTPWMAPEVIQS 180  
 DB 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTTPWMAPEVIQS 180  
 QY 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVLVVEKNERLTIPSSCPSPFAELH 240  
 DB 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVLVVEKNERLTIPSSCPSPFAELH 240  
 QY 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKNSFLHNKAWEWCETIATLERLKLKLERD 300  
 DB 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKNSFLHNKAWEWCETIATLERLKLKLERD 300  
 QY 301 LSFKEQELKERERRLKWEQKLTQOSNTPLL 331  
 DB 301 LSFKEQELKERERRLKWEQKLTQOSNTPLL 331

## RESULT 3

Q9HCC5 PRELIMINARY; PRT; 800 AA.  
 AC Q9HCC5;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE MLTK-alpha.  
 GN MLTK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 SEQUENCE FROM N.A.  
 RX MEDLINE=21264927; PubMed=11042189;  
 RA Gotch I., Adachi M., Nishida E.;  
 RT "Identification and Characterization of a Novel MAP Kinase Kinase Kinase, MLTK.";  
 RL J. Biol. Chem. 276:4276-4286(2001).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AB049733; BAB16444.1; -.  
 DR HSSP; P12931; 1FMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYK; 1.  
 DR SMART; SM00220; S\_TK; 1.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.

KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91188 MW; 2C8593824AB3FADD CRC64;

Query Match 72.7%; Score 331; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWISQDKEVAVKLLKIEKAEILSVL 60  
 DB 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWISQDKEVAVKLLKIEKAEILSVL 60  
 QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYINSNRSEEMDMDHMTWATDVAKGMHY 120  
 DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYINSNRSEEMDMDHMTWATDVAKGMHY 120  
 QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTTPWMAPEVIQS 180  
 DB 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTTPWMAPEVIQS 180  
 QY 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVLVVEKNERLTIPSSCPSPFAELH 240  
 DB 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVLVVEKNERLTIPSSCPSPFAELH 240  
 QY 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKNSFLHNKAWEWCETIATLERLKLKLERD 300  
 DB 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKNSFLHNKAWEWCETIATLERLKLKLERD 300  
 QY 301 LSFKEQELKERERRLKWEQKLTQOSNTPLL 331  
 DB 301 LSFKEQELKERERRLKWEQKLTQOSNTPLL 331

## RESULT 4

Q9NYL2 PRELIMINARY; PRT; 800 AA.  
 AC Q9NYL2;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Mixed lineage kinase.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 SEQUENCE FROM N.A.  
 RX MEDLINE=20384179; PubMed=10924358;  
 RA Liu T.C., Huang C.J., Chu Y.C., Wei C.C., Chou C.C., Chou M.Y.,  
 RA Chou C.K., Yang J.J.;  
 RT "Cloning and expression of ZAK, a mixed lineage kinase-like protein containing a leucine-zipper and a sterile-alpha motif.";  
 RL Biochem. Biophys. Res. Commun. 274:811-816(2000).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AF238255; AAF63490.1; -.  
 DR HSSP; P12931; 1FMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYK; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91264 MW; DA82D7ABB2082F43 CRC64;

Query Match 72.7%; Score 331; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSLGASFVQIKFDLLOFFENC GGSGFGSVYRAKMWISQDKEAVAKKLLIKEAEILSVL 60
Db 1 MSLGASFVQIKFDLLOFFENC GGSGFGSVYRAKMWISQDKEAVAKKLLIKEAEILSVL 60
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYINSNRSEEMDMHMTWATDVAKGMHY 120
Db 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYINSNRSEEMDMHMTWATDVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMWAEVIOQ 180
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMWAEVIOQ 180
QY 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240
Db 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNSFLHNKAEMWRCIEATLERLKLKLERD 300
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNSFLHNKAEMWRCIEATLERLKLKLERD 300
QY 301 LSFKEQELKERERRLKMWEOQLTEQSNTPLL 331
Db 301 LSFKEQELKERERRLKMWEOQLTEQSNTPLL 331

```

## RESULT 5

```

Q9N9E9 PRELIMINARY; PRT; 800 AA.
ID Q9N9E9
AC Q9N9E9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Sterile-alpha motif and leucine zipper containing kinase Azk (Mixed lineage kinase-related kinase MRK-alpha).
DE Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA McNeer J.J., Frima N., Diamond T.E., Dower S.K., Guesdon F.;
RT "Cloning and characterisation of Azk, a mixed lineage kinase containing a sterile-alpha motif."
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=21950776; PubMed=11836244;
RT Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;
RT "MRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in gamma-Radiation-induced Cell Cycle Arrest."
RL J. Biol. Chem. 277:13873-13882(2002).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF251441; AAF65822.1; -.
DR EMBL; AF480461; AAL85891.1; -.
DR HSSP; P12931; 1FMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50011; STYKC; 1.
DR PROSITE; PS50018; PROTEIN_KINASE_ST; 1.
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 800 AA; 91181 MW; B289D836EC52E295 CRC64;

```

Query Match 72.7%; Score 331; DB 4; Length 800;  
Best Local Similarity 100.0%; Pred. No.0;  
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSLGASFVQIKFDLLOFFENC GGSGFGSVYRAKMWISQDKEAVAKKLLIKEAEILSVL 60
Db 1 MSLGASFVQIKFDLLOFFENC GGSGFGSVYRAKMWISQDKEAVAKKLLIKEAEILSVL 60
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYINSNRSEEMDMHMTWATDVAKGMHY 120
Db 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYINSNRSEEMDMHMTWATDVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMWAEVIOQ 180
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMWAEVIOQ 180
QY 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240
Db 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNSFLHNKAEMWRCIEATLERLKLKLERD 300
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNSFLHNKAEMWRCIEATLERLKLKLERD 300
QY 301 LSFKEQELKERERRLKMWEOQLTEQSNTPLL 331
Db 301 LSFKEQELKERERRLKMWEOQLTEQSNTPLL 331

```

## RESULT 6

```

Q9ESL3 PRELIMINARY; PRT; 454 AA.
ID Q9ESL3
AC Q9ESL3;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE MLTK-beta.
DE ZAK OR MLTK.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21264927; PubMed=11042189;
RT Gotch I., Adachi M., Nishida E.;
RT "Identification and characterization of a Novel MAP Kinase Kinase kinase, MLTK."
RT J. Biol. Chem. 276:4276-4286(2001).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049732; BAB16443.1; -.
DR HSSP; P12931; 1FMK.
DR MGD; MGI:1931274; Zak.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKC; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50018; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 454 AA; 51366 MW; 35C2FC0D729D9395 CRC64;

```

Query Match 36.5%; Score 166; DB 11; Length 454;  
Best Local Similarity 99.6%; Pred. No. 3.3e-168;  
Matches 266; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MSLGASFVQIKFDLLOFFENC GGSGFGSVYRAKMWISQDKEAVAKKLLIKEAEILSVL 60

```

|    |     |                                      |                              |     |
|----|-----|--------------------------------------|------------------------------|-----|
| Db | 1   | MSSLGASFVQIKFDDLOFFENCGGSGFGSVYRAKWI | SQDKEXAVKLLKIEKAEILSVL       | 60  |
| Qy | 61  | SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYN   | INSNRSEEMDMHIMTWATDVAKMHY    | 120 |
| Db | 61  | SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYN   | INSNRSEEMDMHIMTWATDVAKMHY    | 120 |
| Qy | 121 | LHMEAPVKVIHRDLKSRNVVIAADGVLKICD      | FGASRFHNHTHMSLVGTFPMAAPEVIO  | 180 |
| Db | 121 | LHMEAPVKVIHRDLKSRNVVIAADGVLKICD      | FGASRFHNHTHMSLVGTFPMAAPEVIO  | 180 |
| Qy | 181 | LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQ    | VAMLVVEKNERLTI PSSCPRSPAEHLH | 240 |
| Db | 181 | LPVSETCDTYSYGVVLWEMLTREVPFKGLEGLQ    | VAMLVVEKNERLTI PSSCPRSPAEHLH | 240 |
| Qy | 241 | QCWEADAKRPSFKQIISILESMSNDT           | 267                          |     |
| Db | 241 | QCWEADAKRPSFKQIISILESMSNDT           | 267                          |     |

| ID | Q9ESL4   | PRELIMINARY; | PRT;                | 802 AA.            |
|----|--|--------------|---------------------|--------------------|
| AC | Q9ESL4;  |              |                     |                    |
| DT | 01-MAR-2001 (TREMBLrel. 16, Created)                               |              |                     |                    |
| DT | 01-MAR-2001 (TREMBLrel. 16, Last sequence update)                  |              |                     |                    |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last annotation update)                |              |                     |                    |
| DE | MLTK alpha.  |              |                     |                    |
| GN | ZAK OR MLTK.   |              |                     |                    |
| OS | Mus musculus (Mouse).  |              |                     |                    |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |              |                     |                    |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |              |                     |                    |
| NC | NCB1_TaxId=10090;  |              |                     |                    |
| RN | [1]  |              |                     |                    |
| RP | SEQUENCE FROM N.A.   |              |                     |                    |
| RX | MEDLINE=21264927; PubMed=11042189;                                 |              |                     |                    |
| RA | Gotoh T.; Adachi M.; Nishida E.;                                   |              |                     |                    |
| RT | "Identification and Characterization of a Novel MAP Kinase Kinase  |              |                     |                    |
| RT | kinase, MLTK.";  |              |                     |                    |
| RL | J. Biol. Chem. 276:4276-4286(2001).                                |              |                     |                    |
| CC | -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  |              |                     |                    |
| CC | EMBL; AB049731; BAB16442.1; .                                      |              |                     |                    |
| DR | HSSP; P12931; 1FMK.  |              |                     |                    |
| DR | MGI; MGI:1931274; Zak.   |              |                     |                    |
| DR | InterPro; IPR000194; ATPase a/bcentre.                             |              |                     |                    |
| DR | InterPro; IPR000719; Euk_pkinase.                                  |              |                     |                    |
| DR | InterPro; IPR001660; SAM.  |              |                     |                    |
| DR | InterPro; IPR002290; Ser thr_pkinase.                              |              |                     |                    |
| DR | InterPro; IPR004040; STY_pkinase.                                  |              |                     |                    |
| DR | InterPro; IPR001245; Tyr_pkinase.                                  |              |                     |                    |
| DR | Pfam; PF00069; pkinase; 1.   |              |                     |                    |
| DR | Pfam; PF00536; SAM; 1.   |              |                     |                    |
| DR | PRINTS; PR00109; TYRKINASE.  |              |                     |                    |
| DR | ProDom; PD000001; Euk_pkinase; 1.                                  |              |                     |                    |
| DR | SMART; SM00454; SAM; 1.  |              |                     |                    |
| DR | SMART; SM00221; STYKc; 1.  |              |                     |                    |
| DR | SMART; SM00220; S_TKc; 1.  |              |                     |                    |
| DR | SMART; SM00219; TYKc; 1.   |              |                     |                    |
| DR | PROSITE; PS00152; ATPASE ALPHA BETA; UNKNOWN_1.                    |              |                     |                    |
| DR | PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.                           |              |                     |                    |
| DR | PROSITE; PS00108; PROTEIN_KINASE_ST; 1.                            |              |                     |                    |
| KW | ATP-binding; Serine/threonine-protein kinase; Transferase.         |              |                     |                    |
| SO | SEQUENCE 802 AA; 91719 MW; D431DF8F312A43CC CRC64;                 |              |                     |                    |
| QY | Query Match  | 36.5%;       | Score 166;          | DB 11; Length 802; |
| Db | Best Local Similarity  | 99.6%;       | Pred. No. 5.6e-168; |                    |
| QY | Matches 266; Conservative  | 0;           | Mismatches 1;       | Indels 0; Gaps     |
| QY | 1 MSSLGASFVQIKEDDLQFFENCGGGSGFSGYVRAKWISQDKEVAVKLLKTEKEAAILSVL 60  |              |                     |                    |
| Db | 1 MSSLGASFVQIKEDDLQFFENCGGGSGFSGYVRAKWISQDKEVAVKLLKTEKEAAILSVL 60  |              |                     |                    |
| QY | 61 SHRNIQFYGVILPEPNYGIIVTEYASLSGLYDIYNSNRSEEMDMHIMTWATDVAKGMY 120  |              |                     |                    |

|    |     |   |     |
|----|-----|---|-----|
| Db | 61  | SHRNIIQFYGVILLEPPNYGIWTEYASLGSLYDYINSNRSEENDMEHIMTWATDVAKGMHY | 120 |
| QY | 121 | LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIQS    | 180 |
| Db | 121 | LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIQS    | 180 |
| QY | 181 | LPVSETCDIYSYGVVLWEMLTREVPFKGLEGLQVAVWLVEKNERLTI PSSCPRSFAELLH | 240 |
| Db | 181 | LPVSETCDIYSYGVVLWEMLTREVPFKGLEGLQVAVWLVEKNERLTI PSSCPRSFAELLH | 240 |
| QY | 241 | QCWEADAKRPPSFKOIISILESMSNDT 267                               |     |
| Db | 241 | QCWEADAKRPPSFKOIISILESMSNDT 267                               |     |

```

RESULT 8
Q90ZY8      PRELIMINARY;          PRT;       371 AA.
ID Q90ZY8
AC Q90ZY8;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Protein kinase Npk.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Chou C.-M., Lee I.-L., Leu J.-H., Huang C.-J.;
RT "A novel protein kinase, zNPK, from the zebrafish.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF265343; AAK52416.1; -.
DR InterPro; IPR000719; Euk_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 371 AA; 42456 MW; 9B918BA8B20D296 CRC64;

Query Match           10.8%; Score 49; DB 13; Length 371;
Best Local Similarity 100.0%; Pred. No. 2.6e-43;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 160 HTTHMSLVGTFPMWAPVETIOSLPVSETCDTYSYGVLWEMLTREVPFKG 208
    |||||||
Db 191 HTTHMSLVGTFPMAPEVIQSILPVSETCDTYSYGVLWEMLTREVPFKG 239

RESULT 9
ID 001700      PRELIMINARY;          PRT;       855 AA.
AC 001700;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical 95.1 kDa protein F33E2.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lennard N.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL022593; CAA18635.2; -.
DR EMBL; Z84574; CAA18635.2; JOINED.
```



DR EMBL; Z84574; CAB06544.2; -. JOINED.  
 DR EMBL; AL022593; CAB06544.2; JOINED.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Hypothetical protein; Serine/threonine-protein kinase;  
 KW Transferase.  
 SQ SEQUENCE 855 AA; 95786 MW; F44D2538CB7D95A CRC64;

Query Match 2.4%; Score 11; DB 5; Length 855;  
 Best Local Similarity 100.0%; Pred. No. 0.021;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

193 GVLWEMLTRE 203  
 |||||  
 DB 235 GVLWEMLTRE 245

## RESULT 10

Q9VW24 PRELIMINARY; PRT; 977 AA.

AC Q9VW24; 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE CG8789 protein (LD14856p).  
 GN CG8789.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,  
 Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 Jalaali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,  
 Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nuno J., Pacle J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celniker S.;

RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AE003516; AAF49129.1; -.  
 DR EMBL; AY094787; AAM11140.1; -.  
 DR HSSP; P08631; 1AD5.  
 DR FlyBase; FBgn0036896; CG8789.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 977 AA; 109567 MW; 5CD235939C14BD4E CRC64;

Query Match 2.4%; Score 11; DB 5; Length 977;  
 Best Local Similarity 100.0%; Pred. No. 0.024;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 191 SYGVLMWMLT 201  
 |||||  
 DB 331 SYGVLMWMLT 341

## RESULT 11

Q9SMA2 PRELIMINARY; PRT; 141 AA.

AC Q9SMA2; 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE Zwh0007.1.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. INDICA;  
 RA Hong G., Zhao W.;  
 RT "Oryza sativa genomic DNA, chromosome 4, clone: b6015";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL117264; CAB55391.1; -.  
 SQ SEQUENCE 141 AA; 15696 MW; 1C608AE0459DD77D CRC64;

Query Match 2.0%; Score 9; DB 10; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 0.56;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 439 DDDDDGEEE 447  
 |||||

Db 26 DDDDDGEEE 34

## RESULT 12

Q8ST12

ID Q8ST12 PRELIMINARY; PRT; 205 AA.

DT 01-JUN-2002 (TReMBLrel. 21, Created)

DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)

DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE Hypothetical protein. 6/101.

OS Dictyostelium discoideum (Slime mold)

OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

OX NCBI\_TaxID=44689;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AX4;

RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,

RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,

RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.,

RA "Sequence and Analysis of Chromosome 2 of Dictyostelium.",

RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

RA EMBL; AC115578; AL92211.1; -

RA Hypothetical protein.

SQ SEQUENCE 205 AA; 22960 MW; 7E3D24AD2290C4D0 CRC64;

## Query Match

Best Local Similarity 2.0%; Score 9; DB 5; Length 205;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDGDDDDDD 443

Db 56 EGDGDDDDDD 64

## RESULT 13

O15318

ID O15318 PRELIMINARY; PRT; 233 AA.

DT 01-JAN-1998 (TReMBLrel. 05, Created)

DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)

DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)

DE RNA polymerase III subunit.

GN RPC32.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=97315201; PubMed=9171375;

RA Wang Z., Roeder R.G.;

RA "Three human RNA polymerase III-specific subunits form a subcomplex

RT with a selective function in specific transcription initiation.";

RL Genes Dev. 11:1315-1326(1997).

DR EMBL; U93868; AAB63676.1; -

SQ SEQUENCE 233 AA; 27299 MW; 16EAB6AFA5002FF CRC64;

## Query Match

Best Local Similarity 2.0%; Score 9; DB 4; Length 233;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDGDDDDDD 443

Db 173 EGDGDDDDDD 181

## RESULT 14

Q98182

ID Q98182 PRELIMINARY; PRT; 297 AA.

DT 01-FEB-1997 (TReMBLrel. 02, Created)

DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE MC011L.

GN MC011L.

OS Molluscum contagiosum virus subtype 1 (MCV1).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Molluscipoxvirus.

OX NCBI\_TaxID=10280;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96325459; PubMed=8670425;

RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,

RA Moss B.;

RT "Genome sequence of a human tumorigenic poxvirus: Prediction of

RT specific host response-evasion genes.";

RL Science 273:813-816(1996).

RN [2]

RP SEQUENCE FROM N.A.

RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,

RA Moss B.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; U60315; AAC55139.1; -

DR InterPro; IPR001230; Prenyl site.

DR PROSITE; PS00294; PRENYLATION; UNKNOWN 1.

SQ SEQUENCE 297 AA; 31500 MW; 861999E36CD8A781 CRC64;

Query Match

Best Local Similarity 2.0%; Score 9; DB 12; Length 297;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDDDG 444

Db 206 GDDDDDDDDG 214

RESULT 15

Q39585

ID Q39585 PRELIMINARY; PRT; 408 AA.

AC Q39585;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE Protein kinase.

OS Chlamydomonas reinhardtii.

OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

OC Chlamydomonadaceae; Chlamydomonas.

OX NCBI\_TaxID=3055;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=21GR(+);

RX MEDLINE=96133873; PubMed=8552645;

RA Kurvari V., Zhang Y., Luo Y., Snell W.J.;

RA "Molecular cloning of a protein kinase whose phosphorylation is

RT regulated by genetic adhesion during Chlamydomonas fertilization.";

RL Proc. Natl. Acad. Sci. U.S.A. 93:39-43(1996).

DR EMBL; U36196; AAA96956.1; -

DR InterPro; IPR000644; CBS\_domain.

DR Pfam; PF00571; CBS; 3.

DR SMART; SM00116; CBS; 3.

KW kinase.

SQ SEQUENCE 408 AA; 44946 MW; DBDE6EB30C49CF2 CRC64;

Query Match

Best Local Similarity 2.0%; Score 9; DB 10; Length 408;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDGDDDDDD 443

Db 393 EGDGDDDDDD 401

RESULT 16

Q8S5D1

ID Q8S5D1 PRELIMINARY; PRT; 453 AA.

AC Q85SD1; 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Hypothetical 49.5 kDa protein.  
GN OJ1341F06.5.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA McCombie W.R., Spiegel L., de la Bastide M., Preston R., Ferraro K.,  
RA Kuit K., Nascimento L., Zutavern T., Balija V., Bell M., Baker J.,  
RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,  
RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;  
RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone  
OJ1341F06, from chromosome 10, complete sequence."  
Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
EMBL; AC116926; AAM0885.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 453 AA; 49540 MW; 3661B92E9239644 CRC64;  
Query Match 2.0%; Score 9; DB 10; Length 453;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDDDDG 444  
Db 277 GDDDDDDDDG 285  
RESULT 17  
Q85SG9 PRELIMINARY; PRT; 460 AA.  
ID Q85SG9;  
AC Q85SG9; 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Hypothetical 50.4 kDa protein.  
GN OSJNB0091J06.10.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,  
RA Kuit K., Nascimento L., Zutavern T., Balija V., Bell M., Baker J.,  
RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,  
RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;  
RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone  
OSJNB0091J06, from chromosome 10, complete sequence."  
Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
EMBL; AC113338; AAM08654.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 460 AA; 50423 MW; 4624B7D90845941A CRC64;  
Query Match 2.0%; Score 9; DB 10; Length 460;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDDDDG 444  
Db 284 GDDDDDDDDG 292  
RESULT 18  
Q91YV1

ID Q91YV1 PRELIMINARY; PRT; 502 AA.  
AC Q91YV1; 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Similar to activin A receptor, type II-like 1.  
GN ACVR1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC014291; AAH14291.1; -  
DR MGI; MGI:1338946; Acvr1l.  
DR InterPro; IPR000472; Activin\_rec.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF01064; Activin\_rec; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
KW ATP-binding; Receptor; Transferase.  
SQ SEQUENCE 502 AA; 56519 MW; 439510D3CC740D65 CRC64;  
Query Match 2.0%; Score 9; DB 11; Length 502;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 131 HRDLKSRNV 139  
Db 327 HRDLKSRNV 335  
RESULT 19  
Q91YR0 PRELIMINARY; PRT; 502 AA.  
ID Q91YR0;  
AC Q91YR0; 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Activin A receptor, type II-like 1.  
GN ACVR1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC015083; AAH15083.1; -  
DR MGI; MGI:1338946; Acvr1l.  
DR InterPro; IPR000472; Activin\_rec.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF01064; Activin\_rec; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
KW ATP-binding; Receptor; Transferase.  
SQ SEQUENCE 502 AA; 56483 MW; 800E8254703AE875 CRC64;  
Query Match 2.0%; Score 9; DB 11; Length 502;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 HRDLKSRNV 139  
Db 327 HRDLKSRNV 335

## RESULT 20

Q63559

ID Q63559 PRELIMINARY; PRT; 504 AA.

AC Q63559; 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)  
DE Serine/threonine kinase receptor.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;  
MEDLINE=96198306; PubMed=8928814;  
KA Panchenko M.P., Williams M.C., Brody J.S., Yu Q.;  
RT "Type I receptor serine-threonine kinase preferentially expressed in  
RT pulmonary blood vessels.";  
RL Am. J. Physiol. 270:L547-L558(1996).  
CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; L36088; AAC37705.1; -;  
DR InterPro; IPR000472; Activin\_rec.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR003605; TGFbeta\_GS.  
DR Pfam; PF01064; Activin\_rec; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00467; GS; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 504 AA; 56703 MW; B40EA30775223C8F CRC64;

Query Match 2.0%; Score 9; DB 11; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 HRDLKSRNV 139  
Db 329 HRDLKSRNV 337

## RESULT 21

Q9FG45

ID Q9FG45 PRELIMINARY; PRT; 607 AA.

AC Q9FG45; 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)  
DE Similarity to unknown protein.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
KA Kaneko T., Kato T., Asamizu E., Sato S., Nakamura Y., Kotani H.,  
RA Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP000607; BAB10964.1; -;  
SQ SEQUENCE 607 AA; 70548 MW; 6A9C3AF08FF94634 CRC64;

Query Match 2.0%; Score 9; DB 10; Length 607;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 436 GDDDDDDG 444  
Db 586 GDDDDDDG 594

## RESULT 22

Q9LKB4

ID Q9LKB4 PRELIMINARY; PRT; 611 AA.

AC Q9LKB4; 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)  
DE Genomic DNA, chromosome 3, TAC clone:K15M2.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
KA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RX MEDLINE=20363099; PubMed=10907853;  
RA Nakamura Y.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.  
RT Sequence features of the regions of 4,251,695 bp covered by ninety p1,  
RT TAC and BAC clones.";  
RL DNA Res. 7:217-221(2000).  
DR EMBL; AP000370; BAA97053.1; -;  
SQ SEQUENCE 611 AA; 71221 MW; ACE12328A9429778 CRC64;

Query Match 2.0%; Score 9; DB 10; Length 611;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDDG 445  
Db 446 DDDDDDDG 454

## RESULT 23

Q9GRH2

ID Q9GRH2 PRELIMINARY; PRT; 616 AA.

AC Q9GRH2; 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)  
DE Tyrosine kinase.  
OS Sycon raphanus.  
OC Eukaryota; Metazoa; Porifera; Calcarea; Calcaronea; Leucosoleniida;  
OC Sycetidae.  
OX NCBI\_TaxID=56443;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Skorokhod A., Gamulin V., Gundacker D., Kavan V., Mueller I.M.,  
RA Mueller W.E.G.;  
RT "Origin of insulin receptor tyrosine kinases in marine sponges.";  
RL Biol. Bull. 197:198-206(2000).  
DR EMBL; Y17877; CAC14729.1; -;  
DR HSSP; P06213; IIRK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002011; RTKinaseII.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.



DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Euk\_Pkinase; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00219; TYRK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
KW ATP-binding; Kinase; Transferase.  
SQ SEQUENCE 616 AA; 69477 MW; D573241383DBE1E4 CRC64;

Query Match 2.0%; Score 9; DB 5; Length 616;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVLWEM 199  
Db 425 SYGVVLWEM 433

ULT 24  
ID 09A0J5 PRELIMINARY; PRT; 652 AA.  
AC 09A0J5;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE DNA ligase (EC 6.5.1.2) (Polydeoxyribonucleotide synthase [NAD+]).  
GN LIG OR SPY0751.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Streptococcaceae; Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;  
RX MEDLINE=21192684; Pubmed=11296296;  
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najaf F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
RA "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";  
Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
RT FUNCTION: THIS PROTEIN CATALYZES THE FORMATION OF PHOSPHODIESTER LINKAGES BETWEEN 5'-PHOSPHORYL AND 3'-HYDROXYL GROUPS IN DOUBLE-STRANDED DNA USING NAD AS A COENZYME AND AS THE ENERGY SOURCE FOR THE REACTION. IT IS ESSENTIAL FOR DNA REPLICATION AND REPAIR OF DAMAGED DNA (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: NAD(+) + (DEOXYRIBONUCLEOTIDE) (N) + (DEOXYRIBONUCLEOTIDE) (M) = AMP + NICOTINAMIDE NUCLEOTIDE + (DEOXYRIBONUCLEOTIDE) (N+M).  
CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT DNA LIGASE FAMILY.  
DR EMBL; AE006527; AAK33695.1; -  
DR HSSP; O87703; 1B04.  
DR InterPro; IPR001357; BRCT.  
DR InterPro; IPR001679; DNALigase.  
DR InterPro; IPR004150; DNA\_ligase\_OB.  
DR InterPro; IPR000445; HhH.  
DR InterPro; IPR003583; HhH.1.  
DR InterPro; IPR004149; znf\_DNALigase\_C4.  
DR InterPro; IPR00533; BRCT.1.  
DR Pfam; PF01653; DNA\_ligase\_N; 1.  
DR Pfam; PF03120; DNA\_ligase\_OB; 1.  
DR Pfam; PF03119; DNA\_ligase\_ZBD; 1.  
DR ProDom; PD003944; DNALigase; 1.  
DR SMART; SM00292; BRCT; 1.  
DR SMART; SM00278; HhH1; 3.  
DR SMART; SM00532; LIGANC; 1.  
DR TIGRFAMs; TIGR00575; dnj; 1.  
DR PROSITE; PS0172; BRCT; 1.  
DR PROSITE; PS01055; DNA\_LIGASE\_N1; 1.  
DR PROSITE; PS01056; DNA\_LIGASE\_N2; 1.

KW DNA repair; DNA replication; Ligase; NAD; Complete proteome.  
SQ SEQUENCE 652 AA; 72420 MW; B88D93C9FD638BE5 CRC64;

Query Match 2.0%; Score 9; DB 16; Length 652;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAIIISV 59  
Db 307 EKEAIIISV 315

RESULT 25

ID 012500 PRELIMINARY; PRT; 764 AA.  
AC 012500;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Chromosome XII reading frame ORF YLR114C.  
GN EFR4 OR L9354.8 OR L2941 OR YLR114C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Verhasselt P., Voet M., Volckaert G.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MIPS;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RA Geisel C.;  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Favello A., Fulton L., Gattung S., Greco T., Kireten J., Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D., Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L., Taich A., Trevasakis E., Vignati D., Wilcox L., Woldman P., Vaudin M., Wilson R., Waterston R.;  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RA Waterston R.;  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z73286; CAA97681.1; -  
DR EMBL; U53878; AAB67559.1; -  
DR EMBL; X89514; CAA61692.1; -  
DR SGD; S0004104; EFR4.  
SQ SEQUENCE 764 AA; 86425 MW; 159A00CF933A5C4A CRC64;

Query Match 2.0%; Score 9; DB 3; Length 764;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGE 445  
Db 628 DDDDDDDGE 636

RESULT 26  
Q9XBP8 PRELIMINARY; PRT; 789 AA.  
ID Q9XBP8  
AC Q9XBP8;

DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Serine/threonine kinase PKX11.  
 GN PKX11.  
 OS Myxococcus xanthus.  
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;  
 CC Myxococcales; Cytoobacterineae; Myxococcaceae; Myxococcus.  
 OX NCBI\_TaxID=34;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DZF1;  
 RA Inouye S., Jain R., Ueki T., Nariya H., Xu C., Hsu M.,  
 Munoz-Dorado J., Farez-Vidal E., Inouye M.;  
 RT "Sequence Analysis of 13 Eukaryotic-like Protein Ser/Thr Kinases of  
 Myxococcus xanthus, a Developmental Bacterium and Significance of  
 Their Coexistence with Protein His Kinases."  
 RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF159690; AAD42854.1; -;  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 789 AA; 81613 MW; EE1E764F0BD77C63 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 789;  
 Best Local Similarity 100.0%; Pred. No. 2.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 193 GVLWLEMLT 201  
 ID |||||  
 Db 210 GVLWLEMLT 218

RESULT 27  
 Q9ZSM8 PRELIMINARY; PRT; 856 AA.  
 AC Q9ZSM8;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE E2A1 (AT490202/T10M13.3).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bilodeau P., Luo M., Dennis E.S., Peacock W.J., Chaudhury A.M.;  
 RT "E2A1, a novel polycomb group gene from Arabidopsis thaliana."  
 RL Plant Physiol. 0:0-0(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,  
 Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
 Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
 Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,  
 Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
 Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,  
 Ecker J.R.;  
 RT "Arabidopsis cDNA clones."  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Shin P., Banh J.,  
 Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,  
 Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,

RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
 Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
 Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,  
 Ecker J.R.;  
 RT "Arabidopsis ORF clones."  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.  
 DR EMBL; AF100163; AAD09108.1; -;  
 DR EMBL; AY057477; AAL09711.1; -;  
 DR EMBL; AY090293; AAL90954.1; -;  
 DR HSSP; P10969; 1WGT.  
 DR InterPro; IPR001005; MYB\_DNA\_binding.  
 DR InterPro; IPR001214; SET.  
 DR InterPro; IPR002221; WAP.  
 DR Pfam; PF00856; SET; 1.  
 DR SMART; SM00395; SANT; 1.  
 DR SMART; SM00317; SET; 1.  
 DR PROSITE; PS00317; 4\_DISULFIDE\_CORE; UNKNOWN\_1.  
 DR PROSITE; PS50280; SET; 1.  
 SQ SEQUENCE 856 AA; 95396 MW; DD4B099C936F197C CRC64;

Query Match 2.0%; Score 9; DB 10; Length 856;  
 Best Local Similarity 100.0%; Pred. No. 2.9;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 439 DDDDDGEEE 447  
 ID |||||  
 Db 17 DDDDDGEEE 25

RESULT 28  
 O04246 PRELIMINARY; PRT; 898 AA.  
 AC O04246;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Putative CURLYLEAF-like 1 homeotic protein (POLYCOMB group-like  
 protein).  
 GN T10M13.3 OR AT4G02020.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Johnson A.F., de la Bastide M., Lodhi M., Hoffman J., Hasegawa A.,  
 Gnoj L., Gottesman T., Granat S., Hameed A., Kaplan N., Schutz K.,  
 Shohdy N., Van Keuren K., Parnell L., Dedhia N., Martienssen R.,  
 McComble W.;  
 RT "The sequence of the Arabidopsis thaliana T10M13 BAC."  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Spiegel L.A., Huang E.N., Nascimeto L.U., de la Bastide M., Vil D.M.,  
 Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,  
 Sheker M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,  
 Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.  
 DR EMBL; AF001308; AAC78694.1; -;  
 DR EMBL; AL161493; CAB80695.1; -;  
 DR HSSP; P10969; 1WGT.  
 DR InterPro; IPR001005; MYB\_DNA\_binding.  
 DR InterPro; IPR001214; SET.  
 DR InterPro; IPR002221; WAP.  
 DR Pfam; PF00856; SET; 1.

```
DR SMART; SM00395; SANT; 1.
DR SMART; SM00317; SET; 1.
DR PROSITE; PS00317; 4_DISULFIDE_CORE; UNKNOWN_1.
DR PROSITE; PS50280; SET; 1.
SQ SEQUENCE 898 AA; 100379 MW; 79BE740689C99C2E CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 10; Length 898;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 439 DDDDDGEEE 447
DB 83 DDDDDGEEE 91

RESULT 29
O9P8G2 PRELIMINARY; PRT; 1059 AA.
AC O9P8G2;
01-OCT-2000 (TREMBlrel. 15, Created)
01-OCT-2000 (TREMBlrel. 15, Last sequence update)
01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Potassium transporter Trk1p.
GN TRK1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RA Miranda M., Bashl E., Slayman C.L.;
RT "Cloning and characterization of a TRK homolog from Candida albicans
RT (TRK1) by complementation in Saccharomyces."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF267125; AAF72203.1;
DR InterPro; IPR003445; Cat_trnsp.
DR InterPro; IPR004773; Ktransp_euk.
DR Pfam; PF02386; TrkH; 1.
DR TIGRFAMs; TIGR00934; 2a38euk; 1.
SQ SEQUENCE 1059 AA; 119788 MW; 9C643485C9453747 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 3; Length 1059;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGE 445
DB 482 DDDDDDDGE 490

RESULT 30
O9U5A8 PRELIMINARY; PRT; 1472 AA.
AC O9U5A8;
01-MAY-2000 (TREMBlrel. 13, Created)
01-MAY-2000 (TREMBlrel. 13, Last sequence update)
01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Insulin receptor-like protein precursor.
GN BIR.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY VITELLOGENIC FOLLICLES;
RA Lindstrom-Dinnetz I., Iatrou K.;
RT "Cloning and functional characterization of an insulin receptor-like
RT mRNA expressed in the silkworm ovary."
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC TYROSINE PHOSPHATE.
```

```
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AF025542; AAF21243.1; -.
DR HSSP; P06213; 1IRK.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR002011; RTKinaseII.
DR InterPro; IPR001245; Tyr_Pkinase.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00261; FU; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
KW ATP-binding; Glycoprotein; phosphorylation; Receptor; Signal;
KW Transferase; Transmembrane; Tyrosine-protein kinase.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 1472 INSULIN RECEPTOR-LIKE PROTEIN.
SQ SEQUENCE 1472 AA; 164582 MW; CB6631C67906AF8E CRC64;
```

```
Query Match
Best Local Similarity 2.0%; Score 9; DB 5; Length 1472;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 191 SYGVVLWEM 199
DB 1283 SYGVVLWEM 1291
```

```
RESULT 31
O8T5H2 PRELIMINARY; PRT; 1978 AA.
ID O8T5H2;
AC O8T5H2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative Tyr/Ser/Thr phosphatase.
GN 22J3.6.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Thomasova D., Ton L.Q., Collins F.H., Kafatos F.C.;
RT "Sequencing and analysis of pent region from Anopheles gambiae
RT chromosome 2R."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ439398; CAD28129.1;
DR SEQUENCE 1978 AA; 211417 MW; E0D9506BB19E35F5 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 5; Length 1978;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 438 DDDDDGEEE 446
DB 1733 DDDDDGEEE 1741

RESULT 32
O27732 PRELIMINARY; PRT; 2391 AA.
ID O27732
```

AC Q27732;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)  
DE Copper amine oxidase (EC 1.4.3.6).  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FCQ27;  
RA Flores M.V.C.;  
RT "Molecular Biology of Pyrimidine Biosynthesis in the Human Malarial Parasite Plasmodium falciparum."  
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FCQ27;  
RA Flores M.V.C., O'Sullivan W.J., Stewart T.S.;  
RT "Characterisation of the Carbamoyl Phosphate Synthetase Gene from Plasmodium falciparum."  
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) + H(2)O(2).  
CC -1- COFACTOR: COPPER AND TOPAQUINONE (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE COPPER/TOPAQUINONE OXIDASE FAMILY.  
DR EMBL; L32150; AAA2952.1; --  
DR HSSP; P00968; 1A9X.  
DR InterPro; IPR000901; CPSase.  
DR InterPro; IPR002474; CPSase\_sm\_chain.  
DR InterPro; IPR001317; CPS\_GATase.  
DR InterPro; IPR000269; CUNH\_oxidase.  
DR InterPro; IPR002048; EF-hand.  
DR InterPro; IPR000991; GATase\_1.  
DR InterPro; IPR004362; MGS\_1like.  
DR Pfam; PF00289; CPSase\_L\_chain; 2.  
DR Pfam; PF02786; CPSase\_L\_D2; 2.  
DR Pfam; PF02787; CPSase\_L\_D3; 1.  
DR Pfam; PF00988; CPSase\_sm\_chain; 1.  
DR Pfam; PF00117; GATase; 1.  
DR Pfam; PF02142; MGS; 1.  
DR PRINTS; PR00098; CPSASE.  
DR PRINTS; PR00099; CPSGATASE.  
DR PRINTS; PR00096; GATASE.  
DR PROSITE; PS01164; COPPER\_AMINE\_OXID\_1; 1.  
DR PROSITE; PS00866; CPSASE\_1; 1.  
DR PROSITE; PS00867; CPSASE\_2; UNKNOWN\_2.  
DR PROSITE; PS00018; EF\_HAND; UNKNOWN\_1.  
DR Copper; Llgase; Oxidoreductase; TPQ.  
SQ SEQUENCE 2391 AA; 275674 MW; 426F3C3E612FEE2 CRC64;

Query Match 2.0%; Score 9; DB 5; Length 2391;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 445  
DB 1786 DDDDDDDG 1794

RESULT 33  
ID Q15452 PRELIMINARY; PRT; 101 AA.  
AC Q15452;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, last annotation update)  
DE Protein-serine/threonine kinase (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=94100173; PubMed=8274451;  
RA Schultz S.J., Nigg E.A.;  
RT "Identification of 21 novel human protein kinases, including 3 members of a family related to the cell cycle regulator nima of Aspergillus nidulans."  
RL Cell Growth Differ. 4:821-830(1993).  
DR EMBL; Z25430; CAA80917.1; --  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Euk\_pkinase; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Kinase; Tyrosine-protein kinase.  
FT NON\_TER 1  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 101 AA; 11485 MW; E584845C18F4FCD8 CRC64;

Query Match 1.8%; Score 8; DB 4; Length 101;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
DB 92 WMAPEVIQ 99

RESULT 34  
ID Q9J872 PRELIMINARY; PRT; 106 AA.  
AC Q9J872;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, last annotation update)  
DE ORF63.  
OS Spodoptera exigua nucleopolyhedrovirus.  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
OX NCBI\_TaxID=10454;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20036646; PubMed=10567663;  
RA Iukel W.F., van Strien E.A., Heldens J.G., Broer R., Zuidema D., Goldbach R.W., Vlask J.M.;  
RT "Sequence and organization of the spodoptera exigua multicapsid nucleopolyhedrovirus genome."  
RL J. Gen. Virol. 80:3289-3304(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Iukel W.F.J., van Strien E.A., Heldens J.G.M., Broer R., Zuidema D., Goldbach R.W., Vlask J.M.;  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF169823; AAF33593.1; --  
SQ SEQUENCE 106 AA; 11780 MW; 599AA3E43F88023 CRC64;

Query Match 1.8%; Score 8; DB 12; Length 106;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
DB 11 GDDDDDD 18

RESULT 35  
ID O55500 PRELIMINARY; PRT; 157 AA.  
AC O55500;  
DT 01-JUN-1998 (TReMBLrel. 06, Created)  
DT 01-JUN-1998 (TReMBLrel. 06, last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)  
DE UL51.



GN UL51.  
OS Human cytomegalovirus (strain Towne).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Cytomegalovirus.  
OX NCBI\_TaxID=10363;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TOWNE;  
RX MEDLINE=98241707; PubMed=9573236;  
RA Krosky P.M., Underwood M.R., Turk S.R., Feng K.W., Jain R.K.,  
Ptak R.G., Western A.C., Biron K.K., Townsend L.B., Drach J.C.;  
RT "Resistance of human cytomegalovirus to benzimidazole ribonucleosides  
maps to two open reading frames: UL89 and UL56.";  
RL J. Virol. 72:4721-4728(1998).  
DR EMBL; AF039234; AAC59474.1; -  
DR InterPro; IPR005208; Herpes\_UL33.  
DR Pfam; PF03581; Herpes\_UL33; 1.  
SQ SEQUENCE 157 AA; 16978 MW; 5999035AC484517D CRC64;

Query Match 1.8%; Score 8; DB 12; Length 157;  
Best Local Similarity 100.0%; Pred. No. 7.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 440 DDDGEEE 447  
Db 12 DDDGEEE 19

## RESULT 36

O9J867 PRELIMINARY; PRT; 161 AA.  
AC O9J867;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE ORF68.  
OS Spodoptera exigua nucleopolyhedrovirus.  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
OX NCBI\_TaxID=10454;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20036646; PubMed=10567663;  
RA IJkel W.F., van Strien E.A., Heldens J.G., Broer R., Zuidema D.,  
RA Goldbach R.W., Vlak J.M.;  
RT "Sequence and organization of the spodoptera exigua multicapsid  
nucleopolyhedrovirus genome.";  
RL J. Gen. Virol. 80:3289-3304(1999).

[12]  
SEQUENCE FROM N.A.  
RA IJkel W.F., van Strien E.A., Heldens J.G.M., Broer R., Zuidema D.,  
RA Goldbach R.W., Vlak J.M.;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF169823; AAF33598.1; -  
DR InterPro; IPR002557; Chitin\_bind\_Pera.  
DR Pfam; PF01607; CBM\_14; 1.  
DR SMART; SM00494; ChitBD2; 1.  
SQ SEQUENCE 161 AA; 18509 MW; AE3EA84AD73A9B73 CRC64;

Query Match 1.8%; Score 8; DB 12; Length 161;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
Db 142 DDDDDDDG 149

## RESULT 37

O91902 PRELIMINARY; PRT; 161 AA.  
ID O91902;  
AC O91902;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE CUN017 hypothetical protein.  
GN CUN017.  
OS Culex nigripalpus baculovirus.  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae.  
OX NCBI\_TaxID=130556;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FLORIDA1997;  
RX MEDLINE=21488685; PubMed=11602755;  
RA Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,  
RA Becnel J.J., Rock D.L., Kutish G.F.;  
RT "Genome Sequence of a Baculovirus Pathogenic for Culex nigripalpus.";  
RL J. Virol. 75:11157-11165(2001).

[12]  
SEQUENCE FROM N.A.  
RP STRAIN=FLORIDA1997;  
RC Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,  
RA Becnel J.J., Rock D.L., Kutish G.F.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF403738; AAK94095.1; -  
DR Hypothetical protein.  
SQ SEQUENCE 161 AA; 18553 MW; 99BA93D3FB8386DD CRC64;

Query Match 1.8%; Score 8; DB 12; Length 161;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
Db 116 DDDDDDDG 123

## RESULT 38

O9HW02 PRELIMINARY; PRT; 162 AA.  
ID O9HW02;  
AC O9HW02;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
DE Hypothetical protein PA4129.  
GN PA4129.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
opportunistic pathogen.";  
RL Nature 406:959-964(2000).  
DR EMBL; AE004829; AAG07516.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 162 AA; 18122 MW; 094CA423A6B87795 CRC64;

Query Match 1.8%; Score 8; DB 16; Length 162;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 329 PLLPLAA 336  
Db 34 PLLPLAA 41

## RESULT 39

Q9P840 PRELIMINARY; PRT; 180 AA.  
 AC Q9P840;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE Essential YAEI protein.  
 GN YAEI.  
 OS Candida albicans (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5476;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA De Backer M.D., Logghe M., Vlaene J., Loonen I., Vandoninck S.,  
 RA de Hoogt R., Nelissen B., Dewaele S., Simons F., Verhasselt P.,  
 RA Contreras R., Luyten W.H.M.L.;  
 RT "A novel method for systematic identification of genes required for  
 growth of Candida albicans."  
 Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 EMBL; AJ390499; CAB77639.1; -  
 SO SEQUENCE 180 AA; 20240 MW; EC98B8E63C1676C7 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 3; Length 180;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
 Db 117 GDDDDDD 124

## RESULT 40

000193 PRELIMINARY; PRT; 183 AA.  
 ID 000193  
 AC 000193;  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Small acidic protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=PLACENTA;  
 MEDLINE=97408516; PubMed=9263035;  
 Gong T.W., Hegeman A.D., Shin J.J., Lindberg K.H., Barald K.F.,  
 RA Lomax M.I.;  
 RT "Novel genes expressed in the chick otocyst during development:  
 RT identification using differential display of RNA."  
 RL Int. J. Dev. Neurosci. 15:585-594 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=BRAIN;  
 RC Strauberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=BRAIN;  
 RC Strauberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=BRAIN;  
 RC Strauberg R.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U51678; AAB99729.1; -  
 DR EMBL; BC016352; AAH16352.1; -  
 DR EMBL; BC007103; AAH07103.1; -  
 DR EMBL; BC020937; AAH20937.1; -

SO SEQUENCE 183 AA; 20332 MW; 509BA7377B699A74 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 4; Length 183;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
 Db 118 GDDDDDD 125

## RESULT 41

Q42341 PRELIMINARY; PRT; 184 AA.  
 ID Q42341  
 AC Q42341;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Serine-threonine protein kinase (Fragment).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=CELL SUSPENSION CULTURE OF ECOTYPE COLUMBIA;  
 RA Cooke R., laudie M., Raynal M., Delseny M.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; F19996; CAA23375.1; -  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Kinase; Tyrosine-protein Kinase.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SO SEQUENCE 184 AA; 20933 MW; EB93761356ACE2E0 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 10; Length 184;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIO 179  
 Db 137 WMAPEVIO 144

## RESULT 42

Q9QAM0 PRELIMINARY; PRT; 194 AA.  
 ID Q9QAM0  
 AC Q9QAM0;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Hypothetical 22.3 kDa protein.  
 OS White spot syndrome virus (WSSV).  
 OC Viruses; unclassified viruses.  
 OX NCBI\_TaxID=92652;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20112891; PubMed=10644828;  
 RA van Hulten M.C., Tsai M.F., Schipper C.A., Lo C.F., Kou G.H.,  
 RA Viak J.M.;  
 RT "Analysis of a genomic segment of white spot syndrome virus of shrimp  
 RT containing ribonucleotide reductase genes and repeat regions."  
 RL J. Gen. Virol. 81:307-316 (2000).  
 DR EMBL; AF099142; AAF18485.1; -  
 KW Hypothetical protein.  
 SO SEQUENCE 194 AA; 22267 MW; 73254CF17FE9678C CRC64;

Query Match 1.8%; Score 8; DB 12; Length 194;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDDDD 443  
DB 61 GDDDDDDDD 68

## RESULT 43

O91E7 PRELIMINARY; PRT; 194 AA.  
AC O91E7;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE ORF94.  
OS White spot syndrome virus (WSSV).  
OC Viruses; unclassified viruses.  
NCBI\_TaxID=92652;

[1]  
SEQUENCE FROM N.A.  
MEDLINE=21342572; PubMed=11448154;  
van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,  
Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;  
"The white spot syndrome virus DNA genome sequence.";  
Virology 286:7-22(2001).

[2]  
SEQUENCE FROM N.A.  
van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,  
Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;  
Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
EMBL: AF369029; AAK77763.1;  
SEQUENCE 194 AA; 22281 MW; 19FE26A0504DCC78 CRC64;

Query Match 1.8%; Score 8; DB 12; Length 194;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDDDD 443  
DB 61 GDDDDDDDD 68

## RESULT 44

O8QTE0 PRELIMINARY; PRT; 194 AA.  
AC O8QTE0;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE WSSV234.  
OS White spot syndrome virus (WSSV).  
OC Viruses; unclassified viruses.  
NCBI\_TaxID=92652;

[1]  
SEQUENCE FROM N.A.  
STRAIN=TAIWAN;  
MEDLINE=20517548; PubMed=11062040;  
Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,  
Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;  
"Identification and characterization of a shrimp white spot syndrome  
virus (WSSV) gene that encodes a novel chimeric polypeptide of  
cellular-type thymidine kinase and thymidylate kinase.";  
Virology 277:100-110(2000).

[2]  
SEQUENCE FROM N.A.  
STRAIN=TAIWAN;  
MEDLINE=21844071; PubMed=11853398;  
Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,  
Lo C.F., Kou G.H.;  
"Identification of a nucleocapsid protein (VP35) gene of shrimp white  
spot syndrome virus and characterization of the motif important for

RT targeting VP35 to the nuclei of transfected insect cells.";  
Virology 293:44-53(2002).

[3]  
SEQUENCE FROM N.A.  
STRAIN=TAIWAN;  
Lo C.-F., Kou G.-H.;  
Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
EMBL: AF440570; AAL89102.1;  
DR EMBL: AF440570; AAL89102.1;  
SEQUENCE 194 AA; 22253 MW; 733FFC0B214CCD3D CRC64;

Query Match 1.8%; Score 8; DB 12; Length 194;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDDDD 443  
DB 61 GDDDDDDDD 68

## RESULT 45

O9C7Y9 PRELIMINARY; PRT; 198 AA.  
AC O9C7Y9;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Hypothetical 21.8 kDa protein (Unknown protein).  
GN T2J15.12.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.  
NCBI\_TaxID=3702;

[1]  
SEQUENCE FROM N.A.  
STRAIN=CV. COLUMBIA;  
MEDLINE=21016719; PubMed=11130712;  
Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
Dunn P., Elgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziali A.,  
Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
"Sequence and analysis of chromosome 1 of the plant Arabidopsis  
thaliana.";  
Nature 408:816-820(2000).

[2]  
SEQUENCE FROM N.A.  
Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,  
Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,  
Ecker J., Theologis A., Davis R.W.;  
Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC051631; AAG51532.1;  
DR EMBL: AF386997; AAK62442.1;  
KW Hypothetical protein.  
SEQUENCE 198 AA; 21847 MW; C0AC434A40FE7818 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 198;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
 Db 101 GDDDDDD 108

## RESULT 46

Q39756  
 ID Q39756 PRELIMINARY; PRT; 204 AA.  
 AC Q39756;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE PKF1 protein (Fragment).  
 GN PKF1.  
 OS Fagus sylvatica (Beechnut).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fagales; Fagaceae; Fagus.  
 NCBI\_TaxID=28930;  
 [1]  
 SEQUENCE FROM N.A.  
 RA Nicolas C., Nicolas G., Rodriguez D.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; X97547; CAA66149.1; -  
 DR HSSP; P08631; IAD5.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KM Serine/threonine-protein Kinase.  
 FT NON TER 1  
 SQ SEQUENCE 204 AA; 23305 MW; B4C70D228B1077B0 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 10; Length 204;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRDLKS 136  
 Db 22 VIHRDLKS 29

## RESULT 47

Q39757  
 ID Q39757 PRELIMINARY; PRT; 206 AA.  
 AC Q39757;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE HVLPST.  
 GN HVLPST.  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticaceae; Hordeum.  
 NCBI\_TaxID=4513;  
 [1]  
 SEQUENCE FROM N.A.  
 RA MEDLINE=99324224; PubMed=10393983;  
 RA Feuille C., Keller B.;  
 RT "High gene density is conserved at syntenic loci of small and large  
 RT grass genomes."  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:8265-8270(1999).  
 DR EMBL; AF108009; AAD46470.1; -  
 SQ SEQUENCE 206 AA; 21440 MW; 3F0BF6622035942 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 206;

Best Local Similarity 100.0%; Pred. No. 9.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 436 GDDDDDD 443  
 Db 184 GDDDDDD 191

## RESULT 48

Q945P0  
 ID Q945P0 PRELIMINARY; PRT; 209 AA.  
 AC Q945P0;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Transcriptional activator FHA1.  
 GN FHA1.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 NCBI\_TaxID=4097;  
 [1]  
 SEQUENCE FROM N.A.  
 RA STRAIN=CV. XANTI;  
 RA Kim M.I., An J.W., Pal H.S.;  
 RT "NtFHA1, transcription activator containing forkhead-associated  
 RT domain."  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF411856; AAL05884.1; -  
 DR InterPro; IPR00253; FHA\_domain.  
 DR Pfam; PF00498; FHA; 1.  
 DR PROSITE; PS50006; FHA\_DOMAIN; 1.  
 SQ SEQUENCE 209 AA; 22601 MW; E3440C7327070176 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 10; Length 209;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 Db 164 DDDDDDDG 171

## RESULT 49

Q28960  
 ID Q28960 PRELIMINARY; PRT; 211 AA.  
 AC Q28960;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein AF1309.  
 GN AF1309.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 OC Archaeoglobaceae; Archaeoglobus.  
 NCBI\_TaxID=2234;  
 [1]  
 SEQUENCE FROM N.A.  
 RA STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RC MEDLINE=98049343; PubMed=9389475;  
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,  
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,  
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-



RT reducing archaeon *Archaeoglobus fulgidus*.  
 RL Nature 390:364-370(1997).  
 DR EMBL; AE001013; AAB89938.1; -.  
 DR TIGR; AF1309; -.  
 DR InterPro; IPR002741; DUF56.  
 DR Pfam; PF01879; DUF56; 1.  
 DR Hypothetical protein; Complete proteome.  
 KW SEQUENCE 211 AA; 22882 MW; EB258FASD12855CA CRC64;  
 SQ

Query Match 1.8%; Score 8; DB 17; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 9.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 PULLPLAA 336  
 |||||  
 Db 174 PULLPLAA 181

## RESULT 50

Q91VH4 PRELIMINARY; PRT; 213 AA.

Q91VH4; 01-DEC-2001 (TREMBlrel. 19, Created).  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Hypothetical 23.6 kDa protein (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BREAST TUMOR;  
 RA Strausberg R.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC014284; AAH14284.1; -.  
 DR InterPro; IPR002038; Osteopontin.  
 DR Pfam; PF00865; Osteopontin; 1.  
 KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 213 AA; 23625 MW; 50CCE5DDE9E91797 CRC64;  
 Query Match 1.8%; Score 8; DB 11; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 9.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 |||||  
 7 DDDDDDDG 14

## RESULT 51

Q8RZR9 PRELIMINARY; PRT; 218 AA.

Q8RZR9; 01-JUN-2002 (TREMBlrel. 21, Created).  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE B1140D12.10 protein.  
 GN B1140D12.10.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_Taxid=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC  
 clone:B1140D12.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003537; BAB86162.1; -.

SQ SEQUENCE 218 AA; 23468 MW; 9B158986E7938F29 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 218;  
 Best Local Similarity 100.0%; Pred. No. 9.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 |||||  
 Db 204 DDDDDDDG 211

## RESULT 52

Q8X4P6 PRELIMINARY; PRT; 229 AA.

Q8X4P6; 01-MAR-2002 (TREMBlrel. 20, Created).  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Putative antiterminal Q of prophage CP-933X (Antitermination  
 protein).  
 GN Z1874 OR ECS1620.  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_Taxid=83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 Grobeck E.J., Davis N.W., Lim A., Dimanalta E.T., Potamouis K.,  
 Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 Welch R.A., Blattner F.R.;  
 RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
 RL Nature 409:529-533(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
 Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RA "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 DR EMBL; AE005330; AAG55971.1; -.  
 DR EMBL; AP002555; BAB35043.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 229 AA; 26402 MW; E6ABF81BE9476667 CRC64;

Query Match 1.8%; Score 8; DB 16; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TLERLKL 297  
 |||||  
 Db 142 TLERLKL 149

## RESULT 53

Q50439 PRELIMINARY; PRT; 237 AA.

Q50439; 01-NOV-1996 (TREMBlrel. 01, Created).  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Tnpr protein.  
 GN Tnpr.  
 OS Mycobacterium smegmatis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1772;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92078082; PubMed=1660454;  
RA Cirillo J.D., Barletta R.G., Bloom B.R., Jacobs W.R. Jr.  
RT "A novel transposon trap for mycobacteria: isolation and  
characterization of IS1096."  
RL J. Bacteriol. 173:7772-7780(1991).  
DR EMBL; M76495; AAA98488.1;  
SQ SEQUENCE 237 AA; 26293 MW; B4A0C4952D5C68EE CRC64;

Query Match 1.8%; Score 8; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 431 FDLSEGGD 438  
DB 119 FDLSEGGD 126

## RESULT 54

Q9AS90 PRELIMINARY; PRT; 246 AA.  
AC Q9AS90;  
DT 01-JUN-2001 (TReMBLrel. 17, Created)  
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE P0028E10.1 protein (P0701D05.18 protein).  
GN P0028E10.1 OR P0701D05.18.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
clone:P0028E10."  
Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
clone:P0701D05."  
Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
RA EMBL; AP002912; BAB39898.1;  
DR EMBL; AP003301; BAB64804.1;  
SQ SEQUENCE 246 AA; 25160 MW; F9B39125F8A5A534 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 246;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443  
DB 179 GDDDDDD 186

## RESULT 55

Q9M1H5 PRELIMINARY; PRT; 250 AA.  
AC Q9M1H5;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE Hypothetical 29.2 kDa protein.  
GN T14E10.90.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Obermaier B., Ottenwaelder B., Duchemin D., Zeitler K., Mewes H.W.,  
RA Rudd S., Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL138656; CAB77572.1;  
KW Hypothetical protein.  
SQ SEQUENCE 250 AA; 29183 MW; E153440480BB5F6 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 250;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 RLKLEERD 300  
DB 54 RLKLEERD 61

## RESULT 56

Q9GRQ9 PRELIMINARY; PRT; 258 AA.  
AC Q9GRQ9;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DE Hypothetical 28.1 kDa protein.  
GN L6520.05.  
OS Leishmania major.  
OC Eukaryota; Euzlenzoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_TaxID=5664;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FRIEDLIN;  
RA Masny D., Purnelle B., Goffeau A., Ivens A.C., Quail M.,  
RA Rajandream M.A., Barrell B.G.;  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FRIEDLIN;  
RX MEDLINE=98146435; PubMed=9477341;  
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
RA Smith D.F.;  
RT "A physical map of the Leishmania major Friedlin genome."  
RL Genome Res. 8:135-145(1998).  
DR EMBL; AL446005; CAC14532.1;  
KW Hypothetical protein.  
SQ SEQUENCE 258 AA; 28116 MW; 52A1CB8119F2B079 CRC64;

Query Match 1.8%; Score 8; DB 5; Length 258;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443  
DB 70 GDDDDDD 77

## RESULT 57

Q9CXN1 PRELIMINARY; PRT; 259 AA.  
AC Q9CXN1;  
DT 01-JUN-2001 (TReMBLrel. 17, Created)  
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE 3110052N05R1k protein.  
GN 3110052N05R1K.

OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=EMBRIONIC HEAD;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
DR EMBL; AK014208; BAB29206.1; -  
DR MGD; MGI:1924237; 3110052N05Rik.  
DR InterPro; IPR001454; Hlgase/hydase.  
DR InterPro; IPR002645; STAS.  
DR Pfam; PF00702; Hydrolase; 1.  
SQ SEQUENCE 259 AA; 28730 MW; C8E34D5F0D6C9A8 CRC64;  
  
Query Match 1.8%; Score 8; DB 11; Length 259;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 291 LERLKLLE 298  
DB 57 LERLKLLE 64  
  
RESULT 58  
O9SISO PRELIMINARY; PRT; 260 AA.  
ID O9SISO;  
O9SISO; 01-MAY-2000 (TREMBLrel. 13, Created)  
01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DT At2g25220 protein.  
DE AT2G25220.  
GN AT2G25220.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RX MEDLINE=20083487; PubMed=10617197;  
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,  
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,  
RA Salzberg S.L., Fraser C.M., Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
thaliana."  
RL Nature 402:761-768(1999).  
RN [2]

RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Lin X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AC007070; AAD23669.1; -  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Prodom; PD000001; Euk\_pkinase; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 260 AA; 28793 MW; 5A469E1CB1596BEC CRC64;  
  
Query Match 1.8%; Score 8; DB 10; Length 260;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 129 VHRDLKS 136  
DB 125 VHRDLKS 132  
  
RESULT 59  
O948F2 PRELIMINARY; PRT; 277 AA.  
ID O948F2;  
AC O948F2;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Putative pol. polyprotein.  
GN OSJNBA0013J21.1.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,  
RA Sasaki C., Henry D., Oates R., Simmons J.;  
RT "Rice Genomic Sequence."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC079843; AAL01161.1; -  
KW POLYPROTEIN.  
SQ SEQUENCE 277 AA; 31970 MW; 29730C9E1C3F2027 CRC64;  
  
Query Match 1.8%; Score 8; DB 10; Length 277;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 437 DDDDDDG 444  
DB 75 DDDDDDG 82  
  
RESULT 60  
O64125 PRELIMINARY; PRT; 279 AA.  
ID O64125;  
AC O64125;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Early T-lymphocyte activator-1 (Fragment).  
GN SPPI OR ETA-1/Op.  
OS Mus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10095;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=95272553; PubMed=7753053;  
RA Ono M., Yamamoto T., Nose M.,  
RT "Allelic difference in the nucleotide sequence of the Eta-1/Op gene  
transcript."  
RL Mol. Immunol. 32:447-448(1995).  
DR EMBL: S78177; AAB34351.2; -.  
DR MGD: MGI:98389; Sspl.  
DR InterPro; IPR02038; Osteopontin.  
DR Pfam; PF00865; Osteopontin; 1.  
DR PRINTS; PR00216; OSTEOPONTIN.  
DR SMART; SM00017; OSTEO; 1.  
DR PROSITE; PS00884; OSTEOPONTIN; 1.  
FT NON\_TER 1  
SQ SEQUENCE 279 AA; 30939 MW; BOEF5855D0AC14A2 CRC64;

Query Match 1.8%; Score 8; DB 11; Length 279;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

437 DDDDDDDG 444  
|||||  
73 DDDDDDDG 80

RESULT 61  
ID 002445 PRELIMINARY; PRT; 292 AA.  
AC 002445;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE LOX6 (Fragment).  
GN LOX6.  
OS Helobdella robusta (Leech).  
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;  
OC Rhynchobdellida; Glossiphoniidae; Helobdella.  
OX NCBI\_TaxID=6412;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98008903; PubMed=9344545;  
RA Kourakis M.J., Master V.A., Lohorst D.K., Nardelli-Haeffliger D.,  
RA Wedeen C.J., Martindale M.O., Shankland M.,  
RT "Conserved anterior boundaries of Hox gene expression in the central  
nervous system of the leech Helobdella."  
RL Dev. Biol. 190:284-300(1997).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
DR EMBL; AF004386; AAB61441.1; -.  
DR HSSP; P02833; 9ANT.  
DR TRANSFAC; T03754; -.  
DR InterPro; IPR001827; Antennapedia.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR000047; HTH\_repressr.  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00025; ANTENNAPEDIA.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRINTS; PR00031; HTHREPRESSR.  
DR ProDom; PD000010; Homeobox; 1.  
DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS00032; ANTENNAPEDIA; 1.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
DR PROSITE; PS50071; HOMEBOX\_2; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON\_TER 1  
SQ SEQUENCE 292 AA; 31973 MW; 66F82BBA65B5D674 CRC64;

Query Match 1.8%; Score 8; DB 5; Length 292;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 438 DDDDDDDG 445  
|||||  
Db 194 DDDDDDDG 201

RESULT 62  
ID 09NWX4 PRELIMINARY; PRT; 293 AA.  
AC 09NWX4;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE CDNA FLJ20549 fis, clone KAT11561.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,  
RA Suzuki Y., Obaashi M., Nishi T., Shibahara T., Tanaka T.,  
RA Nakamura Y., Isoigai T., Sugano S.,  
RT "NEDO human cDNA sequencing project."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK000556; BAA91251.1; -.  
DR HSSP; Q00534; 1B18.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000011; Euk\_pkinase; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW ATP-binding; Tyrosine-protein kinase.  
SQ SEQUENCE 293 AA; 32883 MW; 49A1AB8BC97CB03E CRC64;

Query Match 1.8%; Score 8; DB 4; Length 293;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 WMAPEVIQ 179  
|||||  
Db 62 WMAPEVIQ 69

RESULT 63  
ID 017319 PRELIMINARY; PRT; 297 AA.  
AC 017319;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE NK-2 class homeodomain protein.  
GN CEH-24.  
OS Caenorhabditis briggsae.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6238;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Harfe B., Fire A.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
DR EMBL; AF026057; AAB86603.1; -.  
DR HSSP; P23441; 1FTT.  
DR TRANSFAC; T03376; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR ProDom; PD000010; Homeobox; 1.  
DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
DR PROSITE; PS50071; HOMEBOX\_2; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON\_TER 1  
SQ SEQUENCE 297 AA; 32886 MW; 3FAF5A4886D60755 CRC64;

Query Match 1.8%; Score 8; DB 5; Length 297;  
Best Local Similarity 100.0%; Pred. No. 13;



Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDDDD 443  
|||||  
Db 224 GDDDDDDDD 231

## RESULT 64

064936 PRELIMINARY; PRT; 298 AA.  
AC 064936;  
DT 01-AUG-1998 (TREMBlrel. 07, Created)  
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Zinc finger protein (Fragment).  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;

[1]  
SEQUENCE FROM N.A.

RC STRAIN=LANDSBERG ERECTA;  
RA Bilodeau P., Luo M., Dennis E.S., Peacock W.J., Chaudhury A.M.;  
RT "Cloning of a cDNA encoding a two fingered C2H2 zinc-finger protein  
from Arabidopsis thaliana (Accession No. AF030304) (PGR98-048).";  
RL Plant Physiol. 116:1193-1193(1998).  
DR EMBL; AF030304; AAC09174.1;  
DR InterPro; IPR000822; Znf\_C2H2.  
DR Pfam; PF00096; zf-C2H2; 2.  
DR SMART; SM00355; Znf\_C2H2; 2.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 2.  
DR PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 2.  
KW DNA-binding; Metal-binding; Zinc-finger.  
FT NON TER 1  
SQ SEQUENCE 298 AA; 32954 MW; C3E290E9F4227A2D CRC64;

Query Match 1.8%; Score 8; DB 10; Length 298;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
|||||  
Db 25 DDDDDDDG 32

## RESULT 65

08VB23 PRELIMINARY; PRT; 302 AA.

08VB23;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Wsv178.  
OS White spot syndrome virus (WSSV).  
OC Viruses; unclassified viruses.  
OX NCBI\_TaxID=92652;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21548311; PubMed=11689662;  
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
RT "Complete genome sequence of the shrimp white spot bacilliform  
virus.";  
RL J. Virol. 75:11811-11820(2001).

RN [2]  
RP SEQUENCE FROM N.A.  
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF332093; AAL33182.1; 23C439039EA434A1 CRC64;  
SQ SEQUENCE 302 AA; 35369 MW; 23C439039EA434A1 CRC64;

Query Match 1.8%; Score 8; DB 12; Length 302;  
Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDDDD 443  
|||||  
Db 61 GDDDDDDDD 68

## RESULT 66

09NKD3 PRELIMINARY; PRT; 307 AA.  
ID 09NKD3;  
AC 09NKD3;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE Hypothetical 33.9 kDa protein.  
GN BG:DS05899.6.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;

[1]  
SEQUENCE FROM N.A.

RP STRAIN=BERKELEY;  
RC MEDLINE=99403001; PubMed=10471707;  
RX Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,  
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,  
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,  
RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,  
RA Celniker S., Rubin G.M.;  
RT "An exploration of the sequence of a 2.9-Mb region of the genome of  
Drosophila melanogaster: the Adh region.";  
RL Genetics 153:179-219(1999).

RN [2]  
RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;  
RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,  
RA Butenhoif C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,  
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,  
RA Houston K.A., Hummasti S.R., Katta K., Kearney L., Kim E., Lee B.,  
RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,  
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,  
RA Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,  
RA Zieran L.L., Rubin G.M.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003408; AAF44852.1;  
DR FlyBase; FBgn0028873; BG:DS05899.6.  
KW Hypothetical protein.

SQ SEQUENCE 307 AA; 33948 MW; 24E40352F49F8C2F CRC64;

Query Match 1.8%; Score 8; DB 5; Length 307;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
|||||  
Db 82 DDDDDDDG 89

## RESULT 67

P97827 PRELIMINARY; PRT; 317 AA.

ID P97827;  
AC P97827;  
DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Osteopontin.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;

RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=WISTER;  
 RA Yamamoto S.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WISTER;  
 RA YAMAMOTO S., UEMURA T.;  
 RT "Localization of osteopontin fused to the green fluorescent protein in  
 living osteoblastic cells."  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB001382; BAA19247.1;  
 DR InterPro; IPR002038; Osteopontin.  
 DR Pfam; PF00865; Osteopontin; 1.  
 DR PRINTS; PR00216; OSTEOPONTIN.  
 DR SMART; SM00017; OSTEO; 1.  
 DR PROSITE; PS00884; OSTEOPTNTIN; 1.  
 SQ SEQUENCE 317 AA; 34963 MW; 779A5F00F8862310 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 11; Length 317;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
 Db 88 DDDDDDDG 95

RESULT 68  
 O9PTX7  
 ID O9PTX7 PRELIMINARY; PRT; 318 AA.  
 AC O9PTX7;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE Calreticulin (Fragment).  
 OS Lehteneron reissneri.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
 OC Petromyzontiformes; Petromyzontidae; Lehteneron.  
 OX NCBI\_TaxID=7753;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20063780; PubMed=10594174;  
 RA Kuraku S., Hoshiyama D., Kato K., Suga H., Miyata T.;  
 RT "Monophyly of lampreys and hagfishes supported by nuclear DNA-coded  
 genes."  
 RL J. Mol. Evol. 49:729-735 (1999).  
 DR EMBL; AB025328; BAA88481.1;  
 DR InterPro; IPR001580; Calreticulin.  
 DR InterPro; IPR000886; ER\_target.  
 DR Pfam; PF00262; calreticulin; 1.  
 DR PRINTS; PR00626; CALRETICULIN.  
 DR PRODOM; PD001866; Calreticulin; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
 DR PROSITE; PS00014; ER\_TARGET; UNKNOWN\_1.  
 FT NON TER 1  
 SQ SEQUENCE 318 AA; 36997 MW; C88102EALCAG1506 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 13; Length 318;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443  
 Db 285 GDDDDDD 292

RESULT 69  
 O27264  
 ID O27264 PRELIMINARY; PRT; 318 AA.  
 AC O27264;  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Conserved protein.  
 GN MTH1196.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=187420;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DELTA H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 deltaH: functional analysis and comparative genomics."  
 RL J. Bacteriol. 179:7135-7155 (1997).  
 DR EMBL; AE000888; AAB85685.1;  
 DR InterPro; IPR003801; DUF198.  
 DR Pfam; PF02649; DUF198; 1.  
 DR Complete proteome.  
 KW SEQUENCE 318 AA; 36035 MW; 75BC399BCD1535D9 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 17; Length 318;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 VKLLKIE 51  
 Db 28 VKLLKIE 35

RESULT 70  
 O913E5  
 ID O913E5 PRELIMINARY; PRT; 321 AA.  
 AC O913E5;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE ORF1771 (Fragment).  
 GN ORF1771.  
 OS White spot syndrome virus (WSSV).  
 OC Viruses; unclassified viruses.  
 OX NCBI\_TaxID=92652;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang X., Huang C., Hew C.L.;  
 RT "A gene from shrimp white spot syndrome virus."  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY050248; AAL11022.1;  
 FT NON TER 321  
 SQ SEQUENCE 321 AA; 35834 MW; 9D9083F6A1BE3E08 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 12; Length 321;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443  
 Db 170 GDDDDDD 177

RESULT 71  
 O8VAL2  
 ID O8VAL2 PRELIMINARY; PRT; 321 AA.  
 AC O8VAL2;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE wav390.  
OS White spot syndrome virus (WSSV).  
OC Viruses; unclassified viruses.  
OX NCBI\_TaxID=92652;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21548311; PubMed=11689662;  
RA Yang F., He J., Lin X., Li O., Pan D., Zhang X., Xu X.;  
RT "Complete genome sequence of the shrimp white spot bacilliform virus.";  
RL J. Virol. 75:11811-11820(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX Yang F., He J., Lin X., Li O., Pan D., Zhang X., Xu X.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF332093; AAL33392.1; -  
SQ SEQUENCE 321 AA; 35834 MW; 9D9083F6A1B3E08 CRC64;  
Query Match 1.8%; Score 8; DB 12; Length 321;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDDDD 443  
Db 170 GDDDDDDDD 177  
RESULT 72  
Q91L82 PRELIMINARY; PRT; 321 AA.  
AC Q91L82;  
DT 01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE ORF170 (WSSV449).  
OS White spot syndrome virus (WSSV).  
OC Viruses; unclassified viruses.  
OX NCBI\_TaxID=92652;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21342572; PubMed=11448154;  
RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,  
RT Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlask J.M.;  
RL "The white spot syndrome virus DNA genome sequence.";  
RN [2]  
RP SEQUENCE FROM N.A.  
RX van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,  
RT Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlask J.M.;  
RL Virology 286:7-22(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN=TAIWAN;  
DE MEDLINE=20517548; PubMed=11062040;  
RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,  
RT Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;  
RL "Identification and characterization of a shrimp white spot syndrome virus (WSSV) gene that encodes a novel chimeric polypeptide of cellular-type thymidine kinase and thymidylate kinase.";  
RN [4]  
RP SEQUENCE FROM N.A.  
RX STRAIN=TAIWAN;  
DE MEDLINE=21844071; PubMed=11853398;  
RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,  
RT Lo C.F., Kou G.H.;  
RL "Identification of a nucleocapsid protein (VP35) gene of shrimp white spot syndrome virus and characterization of the motif important for targeting VP35 to the nuclei of transfected insect cells.";  
RN [5]  
RP SEQUENCE FROM N.A.  
RX STRAIN=TAIWAN;

RA Lo C.-F., Kou G.-H.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF369029; AAK77839.1; -  
DR EMBL; AF440570; AAL89317.1; -  
SQ SEQUENCE 321 AA; 35833 MW; C39083F6A2DEB805 CRC64;  
Query Match 1.8%; Score 8; DB 12; Length 321;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDDDD 443  
Db 170 GDDDDDDDD 177  
RESULT 73  
Q8S667 PRELIMINARY; PRT; 325 AA.  
AC Q8S667;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Putative protein kinase.  
GN OSJNBA0034B05.26.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,  
RA Sasaki C., Henry D., Oates R., Simmons J.;  
RT "Rice Genomic Sequence.";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC097446; AAM14697.1; -  
KW Kinase.  
SQ SEQUENCE 325 AA; 35532 MW; 415B37F43743F5D0 CRC64;  
Query Match 1.8%; Score 8; DB 10; Length 325;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 147 LKICDFGA 154  
Db 182 LKICDFGA 189  
RESULT 74  
Q8S672 PRELIMINARY; PRT; 326 AA.  
AC Q8S672;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Putative protein kinase.  
GN OSJNBA0034B05.21.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,  
RA Sasaki C., Henry D., Oates R., Simmons J.;  
RT "Rice Genomic Sequence.";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC097446; AAM14692.1; -  
KW Kinase.  
SQ SEQUENCE 326 AA; 35591 MW; 98F6F2BA6BF0EF36 CRC64;  
Query Match 1.8%; Score 8; DB 10; Length 326;

Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 147 LKICDFGA 154

Db 183 LKICDFGA 190

## RESULT 75

Q8T1F9

ID Q8T1F9

AC Q8T1F9;

DT 01-JUN-2002 (TRENBLREL. 21, Created)

DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)

DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)

DE Hypothetical 38.8 kDa protein.

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyostelidida; Dictyostelium.

OX NCBI\_TaxID=44689;

[1]

SEQUENCE FROM N.A.

STRAIN=AX4;

RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,

RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,

RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.,

RT "Sequence and Analysis of Chromosome 2 of Dictyostelium";

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC116958; AAM08439.1; -.

KM Hypothetical protein.

SQ SEQUENCE 335 AA; 38820 MW; 1B274330A885CFE3 CRC64;

Query Match 1.8%; Score 8; DB 5; Length 335;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444

Db 51 DDDDDDDG 58

Search completed: May 1, 2003, 20:55:30  
Job time : 116 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:38:47 ; Search time 40 Seconds

(without alignments)  
1515.726 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455  
Sequence: 1 MSSIGASFVQIKFDLQFFE.....GDDDDDDGEEEDNDMDNSE 455

Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

rd size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 455   | 100.0       | 455    | 21 | AAAB18657   |
| 2          | 455   | 100.0       | 455    | 21 | AAAY83278   |
| 3          | 455   | 100.0       | 455    | 21 | AAAY84321   |
| 4          | 447   | 98.2        | 473    | 22 | AAAM25322   |
| 5          | 331   | 72.7        | 800    | 22 | AAAB71957   |
| 6          | 331   | 72.7        | 800    | 22 | AAAB65673   |
| 7          | 269   | 59.1        | 349    | 22 | AAAG75571   |
| 8          | 124   | 27.3        | 124    | 22 | AAAB65552   |
| 9          | 118   | 25.9        | 141    | 21 | AAAG03583   |
| 10         | 80    | 17.6        | 92     | 22 | AAU87295    |

|    |    |      |      |    |           |                      |
|----|----|------|------|----|-----------|----------------------|
| 11 | 80 | 17.6 | 92   | 22 | AAU17234  | Novel signal trans   |
| 12 | 56 | 12.3 | 144  | 22 | ABG06092  | Novel human diago    |
| 13 | 14 | 3.1  | 14   | 21 | AAAY83280 | Human survival reg   |
| 14 | 13 | 2.9  | 13   | 22 | AAAB71969 | Human TGF-beta rec   |
| 15 | 11 | 2.4  | 977  | 22 | ABAB71694 | Drosophila melanog   |
| 16 | 10 | 2.2  | 598  | 13 | AAAR20470 | Soluble ysc alpha.   |
| 17 | 10 | 2.2  | 729  | 11 | AAAR05457 | KEX1 amino acid se   |
| 18 | 10 | 2.2  | 764  | 17 | AAAM01897 | Nonsense-mediated    |
| 19 | 10 | 2.2  | 764  | 17 | AAAY98056 | Yeast Nmd2p. Sacc    |
| 20 | 10 | 2.2  | 1089 | 17 | AAAM01896 | Yeast Nmd2p. Invol   |
| 21 | 10 | 2.2  | 1089 | 20 | AAAY05835 | Nonsense-mediated    |
| 22 | 10 | 2.2  | 1089 | 21 | AAAG21315 | Yeast Nmd2p. Sacc    |
| 23 | 9  | 2.0  | 133  | 21 | AAAG21315 | Arabidopsis thalia   |
| 24 | 9  | 2.0  | 156  | 21 | AAAG21314 | Arabidopsis thalia   |
| 25 | 9  | 2.0  | 201  | 21 | AAAG07429 | Arabidopsis thalia   |
| 26 | 9  | 2.0  | 224  | 21 | AAAG07428 | Mouse Activin rece   |
| 27 | 9  | 2.0  | 502  | 15 | AAAR55371 | Human mALK-1 clone   |
| 28 | 9  | 2.0  | 502  | 20 | AAAY33304 | MISR3. Rattus rat    |
| 29 | 9  | 2.0  | 505  | 14 | AAAR41922 | Mullerian inhibiti   |
| 30 | 9  | 2.0  | 505  | 17 | AAAM03761 | Streptococcus poly   |
| 31 | 9  | 2.0  | 648  | 23 | ABP30294  | Streptococcus poly   |
| 32 | 9  | 2.0  | 652  | 23 | ABP27599  | Streptococcus poly   |
| 33 | 9  | 2.0  | 652  | 23 | ABP27600  | Streptococcus poly   |
| 34 | 9  | 2.0  | 656  | 23 | AAE16259  | Human kinase PKI-    |
| 35 | 9  | 2.0  | 909  | 22 | AAU03554  | Human protein kina   |
| 36 | 9  | 2.0  | 2391 | 15 | AAAR55694 | Carbamoyl-phosphat   |
| 37 | 8  | 1.8  | 20   | 22 | AAAB83387 | Lunasin fragment #   |
| 38 | 8  | 1.8  | 20   | 22 | AAAB62621 | Soybean lunasin pe   |
| 39 | 8  | 1.8  | 21   | 22 | AAAB83386 | Lunasin fragment #   |
| 40 | 8  | 1.8  | 21   | 22 | AAAB83393 | Lunasin related pe   |
| 41 | 8  | 1.8  | 21   | 22 | AAAB62620 | Soybean lunasin pe   |
| 42 | 8  | 1.8  | 22   | 22 | AAAB83385 | Lunasin fragment #   |
| 43 | 8  | 1.8  | 22   | 22 | ABBB30607 | Peptide #3258 enco   |
| 44 | 8  | 1.8  | 22   | 22 | ABBB35771 | Peptide #3277 enco   |
| 45 | 8  | 1.8  | 22   | 22 | ABBB21195 | Peptide #3194 enco   |
| 46 | 8  | 1.8  | 22   | 22 | AAAM55580 | Human brain expres   |
| 47 | 8  | 1.8  | 22   | 22 | AAAM68960 | Human bone marrow    |
| 48 | 8  | 1.8  | 22   | 22 | AAAM67787 | Peptide #3221 enco   |
| 49 | 8  | 1.8  | 22   | 22 | AAAM29271 | Peptide #3308 enco   |
| 50 | 8  | 1.8  | 22   | 22 | AAAM04504 | Peptide #3186 enco   |
| 51 | 8  | 1.8  | 22   | 22 | AAAB62619 | Soybean lunasin pe   |
| 52 | 8  | 1.8  | 22   | 23 | ABBG38546 | Human peptide enco   |
| 53 | 8  | 1.8  | 25   | 22 | AAAB83395 | Lunasin related pe   |
| 54 | 8  | 1.8  | 31   | 22 | ABBB29408 | Peptide #2059 enco   |
| 55 | 8  | 1.8  | 31   | 22 | ABBB34590 | Peptide #2096 enco   |
| 56 | 8  | 1.8  | 31   | 22 | ABBB19996 | Protein #1995 enco   |
| 57 | 8  | 1.8  | 31   | 22 | AAAM55379 | Human brain expres   |
| 58 | 8  | 1.8  | 31   | 22 | AAAM67774 | Human bone marrow    |
| 59 | 8  | 1.8  | 31   | 22 | AAAM15583 | Peptide #2017 enco   |
| 60 | 8  | 1.8  | 31   | 22 | AAAM28075 | Peptide #2112 enco   |
| 61 | 8  | 1.8  | 31   | 22 | AAAM03327 | Peptide #2009 enco   |
| 62 | 8  | 1.8  | 31   | 23 | ABGG37322 | Human peptide enco   |
| 63 | 8  | 1.8  | 41   | 22 | AAAB83381 | Lunasin fragment #   |
| 64 | 8  | 1.8  | 41   | 22 | AAAB62615 | Soybean lunasin pe   |
| 65 | 8  | 1.8  | 42   | 22 | AAAB83380 | Lunasin fragment #   |
| 66 | 8  | 1.8  | 42   | 22 | ABBB31264 | Peptide #3915 enco   |
| 67 | 8  | 1.8  | 42   | 22 | ABBB36472 | Peptide #3978 enco   |
| 68 | 8  | 1.8  | 42   | 22 | ABBB21815 | Protein #3814 enco   |
| 69 | 8  | 1.8  | 42   | 22 | AAAM57237 | Human brain expres   |
| 70 | 8  | 1.8  | 42   | 22 | AAAM69636 | Human bone marrow    |
| 71 | 8  | 1.8  | 42   | 22 | AAAM17452 | Peptide #3886 enco   |
| 72 | 8  | 1.8  | 42   | 22 | AAAM29972 | Peptide #4009 enco   |
| 73 | 8  | 1.8  | 42   | 22 | AAAM05129 | Peptide #3811 enco   |
| 74 | 8  | 1.8  | 42   | 22 | AAAB62614 | Soybean lunasin pe   |
| 75 | 8  | 1.8  | 42   | 23 | ABGG39258 | Human peptide enco   |
| 76 | 8  | 1.8  | 43   | 22 | AAAB83391 | Lunasin peptide enco |
| 77 | 8  | 1.8  | 43   | 22 | AAAB62613 | Soybean lunasin po   |
| 78 | 8  | 1.8  | 43   | 22 | AAAB37565 | Soybean 2S albumin   |
| 79 | 8  | 1.8  | 49   | 22 | AAAE00682 | Peptide fragment #   |
| 80 | 8  | 1.8  | 68   | 22 | AAU49655  | Propionibacterium    |
| 81 | 8  | 1.8  | 81   | 21 | AAAG56808 | Arabidopsis thalia   |
| 82 | 8  | 1.8  | 100  | 21 | AAAG02194 | Human secreted pro   |
| 83 | 8  | 1.8  | 112  | 20 | AAAY49163 | Soybean PITSURE pr   |

|     |   |     |     |    |          |                    |
|-----|---|-----|-----|----|----------|--------------------|
| 84  | 8 | 1.8 | 117 | 21 | AAG56807 | Arabidopsis thalia |
| 85  | 8 | 1.8 | 134 | 22 | ABG05567 | Novel human diagno |
| 86  | 8 | 1.8 | 138 | 22 | ABG02692 | Novel human diagno |
| 87  | 8 | 1.8 | 158 | 18 | AAW23419 | Soybean albumin 3  |
| 88  | 8 | 1.8 | 158 | 18 | AAW23420 | Chimeric Soybean a |
| 89  | 8 | 1.8 | 158 | 20 | AAW05723 | Soybean Gm2S-1 sul |
| 90  | 8 | 1.8 | 164 | 21 | AAG19672 | Arabidopsis thalia |
| 91  | 8 | 1.8 | 179 | 19 | AAW61156 | cayA1 protein. C   |
| 92  | 8 | 1.8 | 186 | 21 | AAG19671 | Arabidopsis thalia |
| 93  | 8 | 1.8 | 192 | 21 | AAW19671 | Human disease asso |
| 94  | 8 | 1.8 | 192 | 21 | AAW19671 | Human disease asso |
| 95  | 8 | 1.8 | 192 | 21 | AAW19670 | Arabidopsis thalia |
| 96  | 8 | 1.8 | 226 | 17 | AAW95606 | STR7 (suppressor o |
| 97  | 8 | 1.8 | 236 | 22 | AAE00342 | Rat secreted facto |
| 98  | 8 | 1.8 | 236 | 23 | AAU70148 | Rat secreted facto |
| 99  | 8 | 1.8 | 237 | 14 | AAW30996 | IS1096 tnp R prote |
| 100 | 8 | 1.8 | 251 | 20 | AAW83166 | Rat orphan tyrosin |

ALIGNMENTS

RESULT 1  
AAB18657  
ID AAB18657 standard; Protein; 455 AA.

AC AAB18657;

DT 22-JAN-2001 (first entry)

DE A human regulator of intracellular phosphorylation.

XX Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma;  
KW neurological disorder; Parkinson's disease; demyelinating disease;  
KW meningitis; developmental disorder; neuromuscular disorder; cancer;  
KW myasthenia gravis; cell proliferative disorder; actinic keratosis;  
KW arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;  
KW autoimmune disorder; inflammatory disorder; Addison's disease;  
KW acquired immunodeficiency disease; allergy; diabetes mellitus;  
KW rheumatoid arthritis; microbial infection; trauma.  
XX Homo sapiens.  
OS

|    |               |   |
|----|---------------|---|
| XX | Key           | Location/Qualifiers                               |
| FH | Domain        | 16..257   |
| FT | Modified-site | /note= "eukaryotic protein kinase domain"         |
| FT | Modified-site | 61 /note= "potential phosphorylation site"        |
| FT | Modified-site | 89 /note= "potential phosphorylation site"        |
| FT | Modified-site | 96 /note= "potential phosphorylation site"        |
| FT | Modified-site | 97 /note= "potential glycosylation site"          |
| FT | Binding-site  | 129..141 /note= "protein kinase ATP-binding site" |
| FT | Modified-site | 159 /note= "potential glycosylation site"         |
| FT | Modified-site | 234 /note= "potential phosphorylation site"       |
| FT | Modified-site | 252 /note= "potential phosphorylation site"       |
| FT | Modified-site | 258 /note= "potential phosphorylation site"       |
| FT | Modified-site | 265 /note= "potential glycosylation site"         |
| FT | Modified-site | 268 /note= "potential phosphorylation site"       |
| FT | Region        | 294..322 /note= "potential phosphorylation site"  |
| FT | Modified-site | 302 /note= "leucine zipper"                       |
| FT | Modified-site | 302 /note= "potential phosphorylation site"       |

|    |               |   |
|----|---------------|---|
| FT | Modified-site | /note= "potential phosphorylation site"     |
| FT | Modified-site | 342 /note= "potential phosphorylation site" |
| FT | Modified-site | 343 /note= "potential phosphorylation site" |
| FT | Modified-site | 346 /note= "potential phosphorylation site" |
| FT | Modified-site | 364 /note= "potential phosphorylation site" |
| FT | Modified-site | 409 /note= "potential glycosylation site"   |
| FT | Modified-site | 410 /note= "potential phosphorylation site" |
| FT | Modified-site | 414 /note= "potential phosphorylation site" |
| FT | Modified-site | 415 /note= "potential phosphorylation site" |
| FT | Modified-site | 429 /note= "potential phosphorylation site" |
| FT | Modified-site | 434 /note= "potential phosphorylation site" |

WO200055332-A2.

21-SEP-2000.

17-MAR-2000; 2000WO-US07277.

18-MAR-1999; 99US-0125593.

20-MAY-1999; 99US-0135049.

09-JUL-1999; 99US-0143188.

(INCY-) INCYTE PHARM INC.  
Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;  
Lu DAM, Au-Young J;

WPI: 2000-602121/57.  
N-PSDB; AAA75674.

Novel human intracellular phosphorylation regulator polypeptides and  
polynucleotides for diagnosis, prevention and treatment of  
neurological, cell proliferative and autoimmune/inflammatory disorders

Claim 1; Page 75-76; 96pp; English.

The present sequence represents a human regulator of intracellular  
phosphorylation (HRIP). HRIP is useful for screening agonists and  
antagonists of HRIP polypeptide. HRIP and its agonist or antagonist  
are useful for treating a disease or condition associated with  
decreased or increased expression of functional HRIP. Diseases treated  
or diagnosed include neurological disorders such as stroke, Parkinson's  
disease, demyelinating diseases, bacterial and viral meningitis and  
other developmental disorders of the central nervous system,  
neuromuscular disorders, myasthenia gravis, cell proliferative disorders  
such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer  
including leukaemia, melanoma, myeloma and cancer of the adrenal gland,  
bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/  
inflammatory disorder such as Addison's disease, acquired  
immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,  
rheumatoid arthritis, microbial infection and trauma.

Sequence 455 AA;

Query Match 100.0%; Score 455; DB 21; Length 455;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIRFDDLOFPFENCGGSGFVSYYRAKWMISQDKEVAVKLLKIEKEAELSVL 60  
Db 1 MSSLGASFVQIRFDDLOFPFENCGGSGFVSYYRAKWMISQDKEVAVKLLKIEKEAELSVL 60

```

OY 61 SHRNIIOFYGVILLEPPNYGIVTEYASLSGLYDINSNRSEEMDMDHMTWATDAKGMHY 120
DB 61 SHRNIIOFYGVILLEPPNYGIVTEYASLSGLYDINSNRSEEMDMDHMTWATDAKGMHY 120
OY 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPWWADEVIO 180
DB 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPWWADEVIO 180
OY 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEQVAVLVEKNERLTIPSSCPSPFAELLH 240
DB 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEQVAVLVEKNERLTIPSSCPSPFAELLH 240
OY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIEATLERLKKLERD 300
DB 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIEATLERLKKLERD 300
OY 301 LSPKEQELKERERRLKMEQKLTQSNTPLLPLAARMSEESYFESKTEESNSAEMSCOI 360
DB 301 LSPKEQELKERERRLKMEQKLTQSNTPLLPLAARMSEESYFESKTEESNSAEMSCOI 360
OY 361 TATSNGECHGMNPSLQAMLMGFGLDIFSMNKAGAVHSGMOINMOAKONSSKTSKRGRK 420
DB 361 TATSNGECHGMNPSLQAMLMGFGLDIFSMNKAGAVHSGMOINMOAKONSSKTSKRGRK 420
OY 421 KVNMAIGFSDFDLSGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNMAIGFSDFDLSGDDDDDDGEEEDNDMDNSE 455

RESULT 2
AAY83278
ID AAY83278 standard; Protein; 455 AA.
AC AAY83278;
XX
XX 16-AUG-2000 (first entry)
DE Human survival regulating kinase (SRK).
XX
KW Survival regulating kinase; SRK; Raf; MBP; BAD; MAPKK; MAPKK;
KW HAX-1; protein kinase; autophosphorylation; cell growth; regulation;
KW apoptosis; cell survival; nuclear targeting; tumour; human;
KW autoimmune disease.
XX
OS Homo sapiens.
XX
PN WO200022142-A2.
XX
XX 20-APR-2000.
PF 20-SEP-1999; 99WO-US22008.
XX
XX 13-OCT-1998; 98US-0104088.
PR
XX (ONYX-) ONYX PHARM INC.
PA
XX
PI Ruggieri R, Callow M, Diaz P;
XX
XX WPI; 2000-317994/27.
DR N-PSDB; AA293783.
DR
XX
XX Novel human survival regulating kinase polypeptide for screening agents
PT which modulate biological pathways associated with SRK useful in
PT treating autoimmune diseases, tumors and apoptosis-related disorders
XX
XX Claim 4; Figure 2; 62pp; English.
PS
XX Survival regulating kinases (SRK) are a class of proteins involved in
CC cell signal transduction pathways such as mitogen-activated protein
CC kinase pathways. A protein kinase activity means that the SRK can
CC catalyse a reaction in which a phosphate group is transferred from a
CC phosphate donor to a phosphate acceptor amino acid residue,
CC preferably the hydroxyl side chain of a serine or threonine.

```

```

CC Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase
CC activity is similar to that of a MAPKK such as Raf. has a range of
CC other activities including a cell growth-regulatory activity, a cell
CC survival promoting activity, a HAX-1 binding activity, an apoptosis
CC suppressing activity a MAPKK activation or stimulatory activity, a
CC nuclear targeting activity and a SRK-specific immunogenic activity.
CC SRK is useful for identifying agents which modulate cellular
CC transformations mediated by Ras and SRK and agents that modulate the
CC apoptosis suppression activity of SRK. This information may be useful
CC in the treatment of autoimmune diseases, tumours and apoptosis
CC related disorders.
XX
SQ Sequence 455 AA;
Query Match 100.0%; Score 455; DB 21; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MSSIGASFVQIKFDDLOFENCGGSGSVYRAKMTISQDKEVAVKLKEKAEILSVL 60
DB 1 MSSIGASFVQIKFDDLOFENCGGSGSVYRAKMTISQDKEVAVKLKEKAEILSVL 60
OY 61 SHRNIIOFYGVILLEPPNYGIVTEYASLSGLYDINSNRSEEMDMDHMTWATDAKGMHY 120
DB 61 SHRNIIOFYGVILLEPPNYGIVTEYASLSGLYDINSNRSEEMDMDHMTWATDAKGMHY 120
OY 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPWWADEVIO 180
DB 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPWWADEVIO 180
OY 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEQVAVLVEKNERLTIPSSCPSPFAELLH 240
DB 181 LPVSETCDTYSYGVVLWEMLTREVPFKGLEQVAVLVEKNERLTIPSSCPSPFAELLH 240
OY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIEATLERLKKLERD 300
DB 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIEATLERLKKLERD 300
OY 301 LSPKEQELKERERRLKMEQKLTQSNTPLLPLAARMSEESYFESKTEESNSAEMSCOI 360
DB 301 LSPKEQELKERERRLKMEQKLTQSNTPLLPLAARMSEESYFESKTEESNSAEMSCOI 360
OY 361 TATSNGECHGMNPSLQAMLMGFGLDIFSMNKAGAVHSGMOINMOAKONSSKTSKRGRK 420
DB 361 TATSNGECHGMNPSLQAMLMGFGLDIFSMNKAGAVHSGMOINMOAKONSSKTSKRGRK 420
OY 421 KVNMAIGFSDFDLSGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNMAIGFSDFDLSGDDDDDDGEEEDNDMDNSE 455

RESULT 3
AAY84321
ID AAY84321 standard; Protein; 455 AA.
AC AAY84321;
XX
XX 12-JUL-2000 (first entry)
DE A human cardiovascular system associated protein kinase-2.
XX
XX Human; cardiovascular system associated protein kinase-2; CSAPK-2;
KW signalling pathway; cell growth; cell differentiation; gene mapping;
KW tissue typing; forensic identification; cardiovascular disease;
KW congestive heart failure; transgenic animal.
XX
XX Homo sapiens.
XX
XX WO200014212-A1.
PN
XX 16-MAR-2000.
PD
XX
XX 09-SEP-1999; 99WO-US20631.
PF

```



XX 09-SEP-1998; 98US-0099657.  
PR 29-SEP-1998; 98US-0163115.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Acton S;  
XX  
DR WPI; 2000-271053/23.  
DR N-PSDB; AAZ99726, AAZ99727.  
XX  
PT New nucleic acid encoding cardiovascular system associated protein  
PT kinase, used e.g. for diagnosis, treatment and prevention of  
PT cardiovascular disease  
XX  
PS Claim 2; Fig 2; 163pp; English.  
XX  
CC The present sequence represents a human cardiovascular system associated  
CC protein kinase-2 (CSAPK-2). CSAPK polypeptides are involved in signalling  
CC pathways associated with cell growth and differentiation. The CSAPK  
CC polypeptides and polynucleotides are used to screen for agents that  
CC specifically modulate CSAPK, which are potential therapeutic agents.  
CC They are also used for diagnosis, prognosis or monitoring of  
CC CSAPK-related diseases, gene mapping, tissue typing and forensic  
CC identification, and for treating or preventing disorders associated  
CC with aberrant CSAPK expression or activity, especially cardiovascular  
CC diseases such as congestive heart failure. They can also be used in  
CC pharmacogenomics. The CSAPK polynucleotide may also be used to generate  
CC transgenic animals.  
XX  
SQ Sequence 455 AA;  
  
Query Match 100.0%; Score 455; DB 21; Length 455;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MSSLGASFVQIKFDDLOFFENCGGGSGSVYRAKWISQDKEVAVKLLKIEKEAELSVL 60  
Db 1 MSSLGASFVQIKFDDLOFFENCGGGSGSVYRAKWISQDKEVAVKLLKIEKEAELSVL 60  
  
QY 61 SHRNIQFYGVILEPPNYGIVTEYASLGLDYINSNRSEEMDMHMTWATDVAKGMH 120  
Db 61 SHRNIQFYGVILEPPNYGIVTEYASLGLDYINSNRSEEMDMHMTWATDVAKGMH 120  
  
QY 121 LHMEAPVKVIHRDLKSRNVIAADGLKICDFGASRFHNHTTMSLVGTFPMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVIAADGLKICDFGASRFHNHTTMSLVGTFPMAPEVIQS 180  
  
QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAWLVEKNERLTTPSSCPSPFAELH 240  
Db 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGLQVAWLVEKNERLTTPSSCPSPFAELH 240  
  
QY 241 QCWEADAKKRPSFKQIISLESMSNDTSLPDKCNSFLHNKAEMRCEIEATLERLKLERD 300  
Db 241 QCWEADAKKRPSFKQIISLESMSNDTSLPDKCNSFLHNKAEMRCEIEATLERLKLERD 300  
  
QY 301 LSFKEQELKERERRLKWEOKLTEOSNTPLILPLAARMEESYFESKTEESNSAEMSCQI 360  
Db 301 LSFKEQELKERERRLKWEOKLTEOSNTPLILPLAARMEESYFESKTEESNSAEMSCQI 360  
  
QY 361 TATSNGBEGHGMPSLQAMLMGFCDIFSMNKAGAVHSGMGINMOAKONSSKTSKRKG 420  
Db 361 TATSNGBEGHGMPSLQAMLMGFCDIFSMNKAGAVHSGMGINMOAKONSSKTSKRKG 420  
  
QY 421 KVMNALGFSDFDLSEGGDDDDGHEEDNDMDNSE 455  
Db 421 KVMNALGFSDFDLSEGGDDDDGHEEDNDMDNSE 455  
  
RESULT, 4  
AAM25322  
ID AAM25322 standard; Protein; 473 AA.  
XX

AC AAM25322;  
XX  
DT 16-OCT-2001 (first entry)  
XX  
DE Human protein sequence SEQ ID NO:837.  
XX  
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
KW antibacterial; endocrine; cardiac; central nervous system; virucide;  
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
KW antiagregant; haemostatic; vulnery; antilucer; osteopathic; eczema;  
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cyostatic;  
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
KW neuroprotective; haematopoietic disorder; platelet disorder; asthma;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
KW genetic disease; osteoporosis; severe combined immunodeficiency;  
KW thrombocytopaenia; diabetes; multiple sclerosis; depression;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder.  
XX  
OS Homo sapiens.  
XX  
PN MO200153455-A2.  
XX  
PD 26-JUL-2001.  
XX  
PF 22-DEC-2000; 2000MO-US35017.  
XX  
PR 23-DEC-1999; 99US-0471275.  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI; 2001-457603/49.  
DR N-PSDB; AAH99263.  
XX  
PT Isolated human polynucleotides encoding polypeptides, useful for the  
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -  
XX  
PS Claim 20; Page 191; 1217pp; English.  
XX  
CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
CC AAM25963. The proteins can have activities based on the tissues and  
CC cells they are expressed in, such as: antiinflammatory; antirheumatic;  
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;  
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnery;  
CC antilucer; osteopathic; dermatological; antiallergic; antiasthmatic;  
CC antidiabetic; cyostatic; neuroprotective; antidepressant; nootropic;  
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
CC encoding them can be used in gene therapy, antisense therapy and vaccine  
CC production. The proteins and polynucleotides are useful for screening for  
CC agonists or antagonists of a protein and for the treatment and diagnosis  
CC of disorders associated with the activity of a protein e.g. inflammation,  
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders.  
XX  
SQ Sequence 473 AA;  
  
Query Match 98.2%; Score 447; DB 22; Length 473;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENC GGSGFSVYRAKMSODKEVAVKLLKIEKEAETLSVL 60  
DB 19 MSSLGASFVQIKFDDLOFFENC GGSGFSVYRAKMSODKEVAVKLLKIEKEAETLSVL 78  
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMHIMTWATDVAKGMHY 120  
DB 79 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMHIMTWATDVAKGMHY 138  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLIKICDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
DB 139 LHMEAPVKVIHRDLKSRNVVIAADGVLIKICDFGASRFHNHTHMSLVGTFPMAPEVIQS 198  
QY 181 LPVSETCDTYSYGVVLWEMLTREVPEFKGLEQLQVAMLVVEKNERLTISSCPRSPFAELH 240  
DB 199 LPVSETCDTYSYGVVLWEMLTREVPEFKGLEQLQVAMLVVEKNERLTISSCPRSPFAELH 258  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFHLNKAEMRCEIEATLERLKLKLERD 300  
DB 259 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFHLNKAEMRCEIEATLERLKLKLERD 318  
QY 301 LSFKEQELKERERRLKMEQKLTQESNTPLLLPLAARMSEESYFESKTEESNSAEMSCQI 360  
DB 319 LSFKEQELKERERRLKMEQKLTQESNTPLLLPLAARMSEESYFESKTEESNSAEMSCQI 378  
QY 361 TATSNGEGHGMNPSLOAMLMGFEDIFSMNKAGAVMHSGMGINMOAKONSKTTSKRGRK 420  
DB 379 TATSNGEGHGMNPSLOAMLMGFEDIFSMNKAGAVMHSGMGINMOAKONSKTTSKRGRK 438  
QY 421 KVNMAIGFSDFDLSEGGDDDDDDGEEE 447  
DB 439 KVNMAIGFSDFDLSEGGDDDDDDGEEE 465

RESULT 5  
AAB71957 standard; Protein; 800 AA.

ID AAB71957 standard; Protein; 800 AA.  
AC AAB71957;  
XX 11-MAY-2001 (first entry)  
DE Human TGF-beta receptor encoded by cDNA clone HDP5M48.  
XX Human; antisclerotic; dermatological; immunosuppressive; cytostatic;  
KW antinflammatory; anti-HIV; immunostimulant; cardiant; vascular;  
KW ophthalmological; neuroprotective; nootropic; anticonvulsant; vaccine;  
KW antiparkinsonian; antimicrobial; vulnery; gene therapy; infection;  
KW transforming growth factor; TGF; TGF-beta receptor; immune disorder;  
KW hyperproliferative disorder; cardiovascular disease; angiogenesis;  
KW neurological disorder.

OS Homo sapiens.  
XX WO200112670-A1.  
PN 22-FEB-2001.  
XX 10-AUG-2000, 2000WO-US21736.  
PF 13-AUG-1999, 99US-0148682.  
PR 20-SEP-1999, 99US-0154887.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Ruben SM, NI J;  
XX WPI, 2001-202858/20.  
DR N-PSDB; AAF75336.  
XX Nucleic acid molecules encoding 12 transforming growth factor-beta  
PT receptor polypeptides, useful for preventing, diagnosing and treating  
PT e.g. cancers, Parkinson's disease and diabetic retinopathy -

XX Claim 11; Page 293-295; 311pp; English.  
PS The present sequence is one of 12 novel human transforming growth factor  
XX (TGF)-beta receptor polypeptides. The TGF-beta receptor polynucleotides  
CC and polypeptides may be used in the prevention, diagnosis and treatment  
CC of diseases associated with inappropriate polypeptide expression. Such  
CC diseases include immune disorders (e.g. multiple sclerosis, systemic  
CC lupus erythematosus and human immuno-deficiency virus (HIV) infections),  
CC hyperproliferative disorders (e.g. cancers and Gaucher's disease),  
CC cardiovascular diseases (e.g. Scimitar syndrome, Chaga's cardiomyopathy  
CC and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft  
CC neovascularisation and diabetic retinopathy), neurological disorders  
CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease)  
CC and infectious diseases. The polynucleotides and polypeptides are also  
CC useful for promoting wound healing, regeneration and/or chemotaxis. The  
CC polynucleotides and their complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar nucleic acid sequences in samples. The polypeptides may be used  
CC as antigens in the production of antibodies and in assays to identify  
CC modulators of protein expression and activity. The anti-TGF-beta receptor  
CC antibodies may be used to down regulate expression and activity and as  
CC diagnostic agents for detecting the presence of the polypeptides in  
CC samples.  
XX  
SQ Sequence 800 AA;

Query Match 72.7%; Score 331; DB 22; Length 800;  
Best Local Similarity 100.0%; Pred. No. 2.6e-304;  
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENC GGSGFSVYRAKMSODKEVAVKLLKIEKEAETLSVL 60  
DB 1 MSSLGASFVQIKFDDLOFFENC GGSGFSVYRAKMSODKEVAVKLLKIEKEAETLSVL 60  
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMHIMTWATDVAKGMHY 120  
DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMHIMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLIKICDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
DB 121 LHMEAPVKVIHRDLKSRNVVIAADGVLIKICDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
QY 181 LPVSETCDTYSYGVVLWEMLTREVPEFKGLEQLQVAMLVVEKNERLTISSCPRSPFAELH 240  
DB 181 LPVSETCDTYSYGVVLWEMLTREVPEFKGLEQLQVAMLVVEKNERLTISSCPRSPFAELH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFHLNKAEMRCEIEATLERLKLKLERD 300  
DB 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFHLNKAEMRCEIEATLERLKLKLERD 300  
QY 301 LSFKEQELKERERRLKMEQKLTQESNTPLL 331  
DB 301 LSFKEQELKERERRLKMEQKLTQESNTPLL 331

RESULT 6  
AAB65673 standard; Protein; 800 AA.

ID AAB65673 standard; Protein; 800 AA.  
AC AAB65673;  
XX 27-MAR-2001 (first entry)  
DT Novel protein kinase, SEQ ID NO: 201.  
XX Human; mouse; protein kinase; antiarthritis; antisclerotic; osteopathic;  
KW immunosuppressive; cardiant; renal; antinflammatory; antiaschmatic;  
KW dermatological; cardiababetic; antifertility; gene therapy; vaccine;  
KW immune disorder; cardiovascular disease; neurodegenerative disease;  
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.

OS Homo sapiens.  
XX WO200073469-A2.  
XX  
XX  
PD 07-DEC-2000.  
XX  
XX 26-MAY-2000; 2000WO-US14842.  
XX  
XX  
XX 28-MAY-1999; 99US-0136503.  
XX  
XX  
XX (SUGEN-) SUGEN INC.  
XX  
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;  
XX  
XX WPI; 2001-032161/04.  
XX  
XX N-PSDB; AAF44701.  
XX  
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
XX treating immune-related diseases and disorders, cardiovascular disease,  
XX neurodegenerative diseases and/or cancers -  
XX  
XX Claim 10; Fig 1; 310pp; English.

XX The present sequence is a novel protein kinase. The novel protein kinases  
XX and the nucleic acids that encode them may be used in the treatment and  
XX diagnosis of diseases associated with inappropriate kinase expression  
XX such as immune-related diseases and disorders, cardiovascular disease,  
XX neurodegenerative diseases and/or cancers. The nucleic acids and  
XX complementary sequences may also be used as DNA probes in diagnostic  
XX assays. The kinase polypeptides may be used as antigens in the production  
XX of antibodies of kinase expression and activity. Anti-kinase antibodies  
XX and kinase antagonists may also be used to down regulate kinase  
XX expression and activity. Diseases related to kinase expression and  
XX activity include rheumatoid arthritis, atherosclerosis, autoimmune  
XX disorders, complications of organ transplantation, myocardial infarction,  
XX immune disorders, cardiomyopathies, strokes, renal failure,  
XX oxidative-stress related disorders, chronic inflammatory bowel disease,  
XX chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
XX osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
XX reproductive disorders.

XX Sequence 800 AA;

Query Match 72.7%; Score 331; DB 22; Length 800;  
Best Local Similarity 100.0%; Pred. No. 2.6e-304; Indels 0; Gaps 0;  
Matches 331; Conservative 0; Mismatches 0;

QY 1 MSSLGASFVQIKFDLQFFENCSSGSGSVYRAKWIISODKEVAVKLLKIEKAEILSVL 60  
1 MSSLGASFVQIKFDLQFFENCSSGSGSVYRAKWIISODKEVAVKLLKIEKAEILSVL 60  
QY 61 SHRNIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEMDMDHIMTWATDVAKGMHY 120  
61 SHRNIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEMDMDHIMTWATDVAKGMHY 120  
DB 61 SHRNIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEMDMDHIMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLIKICDFGASRFHNHTTMSLVGTFPWWAPEVIQS 180  
121 LHMEAPVKVIHRDLKSRNVIAADGVLIKICDFGASRFHNHTTMSLVGTFPWWAPEVIQS 180  
DB 121 LHMEAPVKVIHRDLKSRNVIAADGVLIKICDFGASRFHNHTTMSLVGTFPWWAPEVIQS 180  
QY 181 LPVSETCDTYSYGVVLMEMLTREVPEKGLQVAMLVVEKNERLTISSCPSPFAELH 240  
181 LPVSETCDTYSYGVVLMEMLTREVPEKGLQVAMLVVEKNERLTISSCPSPFAELH 240  
DB 181 LPVSETCDTYSYGVVLMEMLTREVPEKGLQVAMLVVEKNERLTISSCPSPFAELH 240  
QY 241 QCWEADAKKRPSPKQIISLESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKLKLRD 300  
241 QCWEADAKKRPSPKQIISLESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKLKLRD 300  
DB 241 QCWEADAKKRPSPKQIISLESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKLKLRD 300  
QY 301 LSFKEQELKERERRLKMEQKLTEQSNTPPL 331  
301 LSFKEQELKERERRLKMEQKLTEQSNTPPL 331  
DB 301 LSFKEQELKERERRLKMEQKLTEQSNTPPL 331

RESULT 7

AAG75571  
ID AAG75571 standard; Protein; 349 AA.  
XX  
XX AAG75571;  
AC  
XX  
XX 03-SEP-2001 (first entry)  
DT  
XX  
XX Human colon cancer antigen protein SEQ ID NO:6335.  
DE  
XX  
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200122920-A2.  
PN  
XX  
XX 05-APR-2001.  
PD  
XX  
XX 28-SEP-2000; 2000WO-US26524.  
PF  
XX  
XX 29-SEP-1999; 99US-0157137.  
PR  
XX  
XX 03-NOV-1999; 99US-0163280.  
PR  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX  
XX WPI; 2001-235357/24.  
XX  
XX N-PSDB; AAH34976.  
DR  
XX  
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
XX useful for preventing, diagnosing and/or treating colorectal cancers -  
XX  
XX Claim 11; Page 7789-7790; 9803pp; English.

XX AAH32943 to AAH37195 and AAG77788 represent human colon  
XX cancer-associated nucleic acid molecules (N) and proteins (P), where  
XX the proteins are collectively known as colon cancer antigens. The colon  
XX cancer antigens have cytostatic activity and can be used in gene  
XX therapy and vaccine production. N and P may be used in the prevention,  
XX diagnosis and treatment of diseases associated with inappropriate P  
XX expression. For example, N and P may be used to treat disorders  
XX associated with decreased expression by rectifying mutations or deletions  
XX in a patient's genome that affect the activity of P by expressing P.  
XX Inactive proteins or to supplement the patient's own production of P.  
XX Additionally, N may be used to produce the colon cancer-associated P,  
XX by inserting the nucleic acids into a host cell and culturing the cell  
XX to express the proteins. N and P can be used in the prevention, diagnosis  
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
XX and AAB77789 represent sequences used in the exemplification of the  
XX present invention.  
XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
XX missing at time of publication, meaning no sequences are present for  
XX SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 349 AA;

Query Match 59.1%; Score 269; DB 22; Length 349;  
Best Local Similarity 100.0%; Pred. No. 5.9e-246; Indels 0; Gaps 0;  
Matches 269; Conservative 0; Mismatches 0;

QY 1 MSSLGASFVQIKFDLQFFENCSSGSGSVYRAKWIISODKEVAVKLLKIEKAEILSVL 60  
1 MSSLGASFVQIKFDLQFFENCSSGSGSVYRAKWIISODKEVAVKLLKIEKAEILSVL 60  
DB 57 MSSLGASFVQIKFDLQFFENCSSGSGSVYRAKWIISODKEVAVKLLKIEKAEILSVL 116  
QY 61 SHRNIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEMDMDHIMTWATDVAKGMHY 120  
61 SHRNIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEMDMDHIMTWATDVAKGMHY 120  
DB 61 SHRNIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEMDMDHIMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLIKICDFGASRFHNHTTMSLVGTFPWWAPEVIQS 180  
121 LHMEAPVKVIHRDLKSRNVIAADGVLIKICDFGASRFHNHTTMSLVGTFPWWAPEVIQS 180  
DB 121 LHMEAPVKVIHRDLKSRNVIAADGVLIKICDFGASRFHNHTTMSLVGTFPWWAPEVIQS 180  
QY 177 LHMEAPVKVIHRDLKSRNVIAADGVLIKICDFGASRFHNHTTMSLVGTFPWWAPEVIQS 236  
177 LHMEAPVKVIHRDLKSRNVIAADGVLIKICDFGASRFHNHTTMSLVGTFPWWAPEVIQS 236



|    |     |  |     |
|----|-----|--|-----|
| QY | 181 | LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTIIPSSCPRSFAELIH | 240 |
| Db | 237 | LPVSETCDTYSYGVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTIIPSSCPRSFAELIH | 296 |
| QY | 241 | QCWEADAKKRPSFKOIISILEMSNDTSL                                 | 269 |
| Db | 297 | QCWEADAKKRPSFKOIISILEMSNDTSL                                 | 325 |

|             |                           |
|-------------|---------------------------|
| RESULT 8    | *                         |
| AAB65552    |                           |
| ID AAB65552 | Standard; Protein; 124 AA |
| vv          |                           |

AC AAB65552;

DT 27-MAR-2001 (first entry)

**C-terminus specific to novel human protein kinase MLK4B.**

Human; protein kinase; antiarthritic; antisclerotic; immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic; osteopathic; dermatological; antidiabetic; antifertility; gene therapy; vaccine; immune disorder; cardiovascular disease; neurodegenerative disease; cancer; autoimmune disorder; stroke; inflammatory bowel disease; inflammatory pelvic disease; multiple sclerosis; psoriasis.

**Homo sapiens.**

PN WO200073469-A2.

PD 07-DEC-2000.

26-MAY-2000; 2000WO-US14842.

PR 28-MAY-1999; 99US-0136503.

PA (SUGG-) SUGEN INC.

Plowman GD, Martinez R, Whyte D, Sudersanam S,

WPI; 2001-032161/04.

PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
PT treating immune-related diseases and disorders, cardiovascular disease,  
PT neurodegenerative diseases and/or cancers -  
XX  
PS  
PS Example 1; Page 116; 310pp; English.

PS Example 1; Page 116; 310pp; English.

The present sequence is given in a specification relating to novel protein kinases. The protein kinases and the nucleic acids that encode them may be used in the treatment and diagnosis of diseases associated with inappropriate kinase expression such as immune-related diseases and disorders, cardiovascular disease, neurodegenerative diseases and/or cancers. The nucleic acids and complementary sequences may also be used as DNA probes in diagnostic assays. The kinase polypeptides may be used as antigens in the production of antibodies of kinase expression and activity. Anti-kinase antibodies and kinase antagonists may also be used to down regulate kinase expression and activity. Diseases related to kinase expression and activity include rheumatoid arthritis, atherosclerosis, autoimmune disorders, complications of organ transplantation, myocardial infarction, immune disorders, cardiomyopathies, strokes, renal failure, oxidative-stress related disorders, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive disorders.

Sequence 124 AA;

|                       |                 |                     |           |             |
|-----------------------|-----------------|---------------------|-----------|-------------|
| Query Match           | 27.3%;          | Score 124;          | DB 22;    | Length 124; |
| Best Local Similarity | 100.0%;         | Pred. No. 3.9e-109; |           |             |
| Matches 124;          | Conservative 0; | Mismatches 0;       | Indels 0. | Gaps 0.     |

[illegible]

|    |     |      |     |
|----|-----|------|-----|
| QY | 452 | DNSE | 455 |
| Db | 121 | DNSE | 124 |

RESULT 9  
AAG03583

| ID | standard; protein; 141 AA. |
|----|----------------------------|
| 1  |                            |

AC AAG03583 ;

DT 06-OCT-2000 (first entry)

Human secreted protein, SEQ ID NO: 7664.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping.

05 Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PH 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GEST ) GENSET. ....

PI Dumas Milne Edwards J, Duclert A, Giordano J;  
yy

DR WPI; 2000-500381/45.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures  
XX  
PS Claim 13; SEQ ID 7664; 71pp + CD-ROM; English.  
wv

PS Claim 13; SEQ ID 7664; 71pp + CD-ROM; English

The present sequence is a polypeptide encoded by one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA-sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

**SQ Sequence 141 AA;**

|                       |                 |                   |           |             |
|-----------------------|-----------------|-------------------|-----------|-------------|
| Query Match           | 25.9%;          | Score 118;        | DB 21;    | Length 141; |
| Best Local Similarity | 100.0%;         | Pred. No. 2e-103; |           |             |
| Matches 118;          | Conservative 0; | Mismatches 0;     | Indels 0; | Gaps 0;     |

QY 102 MDMDMHTWATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIADGVLKICDFGASRFHNHT 161

Dp 1 MDMDMHTWATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIADGVLKICDFGASRFHNHT 60



QY 162 THMSLVGTFPMMAPEVIQSLPVSETCDTYSYGVLMEMLTREVPFKGLEGLQVAMLVV 219  
DB 61 THMSLVGTFPMMAPEVIQSLPVSETCDTYSYGVLMEMLTREVPFKGLEGLQVAMLVV 118  
RESULT 10  
AAU87295 \*  
ID AAU87295 standard; Protein; 92 AA.  
XX  
AC AAU87295;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE Novel central nervous system protein #205.  
XX  
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;  
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
KW myocardial infarction; wound healing; cell proliferation; skin aging;  
KW food additive; food preservative; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200155318-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01332.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.

PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249220.  
PR 17-NOV-2000; 2000US-0249221.  
PR 17-NOV-2000; 2000US-0249222.  
PR 17-NOV-2000; 2000US-0249223.  
PR 17-NOV-2000; 2000US-0249224.  
PR 17-NOV-2000; 2000US-0249225.  
PR 17-NOV-2000; 2000US-0249226.  
PR 17-NOV-2000; 2000US-0249227.  
PR 17-NOV-2000; 2000US-0249228.  
PR 17-NOV-2000; 2000US-0249229.  
PR 17-NOV-2000; 2000US-0249230.  
PR 17-NOV-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barab SC, Ruben SM;  
XX  
DR WPI; 2001-581633/65.  
DR N-PSDB; ABK43625.  
XX  
XX

PT New isolated nucleic acid encoding a protein for diagnosing,  
PT preventing, treating or ameliorating medical conditions and used as  
PT food additives or preservatives -  
XX  
XX  
PS Claim 9; SEQ ID No 813; 837bp; English.

XX  
CC The invention describes an isolated nucleic acid molecule (I) encoding a  
CC novel central nervous system protein. (I) and polypeptides (II) encoded  
CC by (I), are used to treat a medical conditions and in diagnosis of a  
CC pathological condition. Disorders which are diagnosed or treated include  
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
CC angiotrophic lateral sclerosis, infections caused by bacteria, viruses  
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders  
CC adenocarcinomas and irritable bowel syndrome, reproductive system  
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes  
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.  
CC leukaemia, disorders involving neovascularisation e.g. malignancies,  
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.  
CC acute kidney failure and blood related disorders e.g. myocardial  
CC infarction. The polypeptides can also be used to aid wound healing and  
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
CC maintain organs before transplantation, for supporting cell culture of  
CC primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 17.6%; Score 80; DB 22; Length 92;  
Best Local Similarity 100.0%; Pred. No. 1.3e-67;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 AEILSVLSHRNIQFYGVLEPPNYGIVTEYASLSGLYDINSNRSEMDMDHMTWATD 113  
DB 13 AEILSVLSHRNIQFYGVLEPPNYGIVTEYASLSGLYDINSNRSEMDMDHMTWATD 72  
QY 114 VAKGMHYLHMEAPVKVTHRD 133  
DB 73 VAKGMHYLHMEAPVKVTHRD 92

RESULT 11  
ID AAU17234  
AC AAU17234; standard; Protein; 92 AA.  
DT 07-NOV-2001 (first entry)  
DE Novel signal transduction pathway protein, Seq ID 799.  
XX  
KW Neuroprotective; cytoskeletal; dermatological; immunosuppressive; tumour;  
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS;  
KW acquired immune deficiency syndrome.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200154733-A1.  
XX  
XX PD 02-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US01312.  
XX

PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 14-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0228287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.

PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-465460/50.  
N-PSDB; AAS27151.

Novel polypeptides useful for diagnosing, treating, preventing and/or  
prognosing disorders related to the proteins, including cancers, immune  
disorders and neuronal disorders -  
Claim 1; SEQ ID No 799; 880pp; English.

The invention relates to novel isolated polypeptides (I), and  
polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
diagnosing, preventing and treating diseases including immune system  
disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
transplant rejection and graft versus host disease, infectious diseases  
(e.g. hepatitis C), bleeding disorders (sickle cell anaemia), myeloproliferative  
other blood-related disorders (leukemia, myelodysplastic syndromes and  
disorders, primary haematopoietic disorders, hyperproliferative  
disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
disorders (e.g. glomerulonephritis), cardiovascular disorders  
(e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
Addison's disease), reproductive system disorders, gastrointestinal  
disorder (inflammatory disorders), liver disorders (cirrhosis),  
as stimulators of B-cell responsiveness to pathogens, activators of  
T-cells, to induce higher affinity antibodies, and as a means to induce  
tumour proliferation in pathologies e.g. acquired immune deficiency  
syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction  
pathway protein, amino acid sequences of the invention.

Query Match 17.6%; Score 80; DB 22; Length 92;  
Best Local Similarity 100.0%; Pred. No. 1.3e-67;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 AEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLGSLDYINSNRSEEMDMHMTWATD 113  
Db 13 AEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLGSLDYINSNRSEEMDMHMTWATD 72  
OY 114 VAKGMHYLMEAPVKVIHRD 133  
Db 73 VAKGMHYLMEAPVKVIHRD 92

RESULT 12

ABG06092  
ID ABG06092 standard; Protein; 144 AA.  
XX  
AC ABG06092;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #6083.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KM food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PA 31-MAR-2000; 2000US-0540217.  
XX 23-AUG-2000; 2000US-0649167.  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
DR N-PSDB; AAS70279.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders of other traits and to assess  
PT biodiversity  
XX  
PS Claim 20; SEQ ID No 36451; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 144 AA;  
QY Query Match 12.3%; Score 56; DB 22; Length 144;  
Best Local Similarity 100.0%; Pred. No. 9.7e-45;  
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 139 VVTAADGVLKICDFGASRFHNTHTMSLVGTFFWMAPEVIOQLPVSETCDTYSYGV 194  
58 VVTAADGVLKICDFGASRFHNTHTMSLVGTFFWMAPEVIOQLPVSETCDTYSYGV 113  
RESULT 13 \*  
ID AAY83280 standard; Protein; 14 AA.  
XX

XX  
AC AAY83280;  
XX  
DT 16-AUG-2000 (first entry)  
XX  
DE Human survival regulating kinase (SRK) antigenic peptide.  
XX  
KW Survival regulating kinase; SRK; Raf; MBP; BAD; MAPKK; MAPKK;  
KM HAX-1; protein kinase; autophosphorylation; cell growth; regulation;  
KM apoptosis; cell survival; nuclear targeting; tumour; human;  
XX autoimmune disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200022142-A2.  
XX  
PD 20-APR-2000.  
XX  
PF 20-SEP-1999; 99WO-US22008.  
XX  
PR 13-OCT-1998; 98US-0104088.  
XX  
PA (ONYX-) ONYX PHARM INC.  
XX  
PI Ruggieri R, Callow M, Diaz P;  
XX  
DR WPI; 2000-317994/27.  
XX  
PT Novel human survival regulating kinase polypeptide for screening agents  
PT which modulate biological pathways associated with SRK useful in  
PT treating autoimmune diseases, tumors and apoptosis-related disorders  
XX  
PS Claim 53; Page 6; 62pp; English.  
XX  
CC Survival regulating kinases (SRK) are a class of proteins involved in  
CC cell signal transduction pathways such as mitogen-activated protein  
CC kinase pathways. A protein kinase activity means that the SRK can  
CC catalyse a reaction in which a phosphate group is transferred from a  
CC phosphate donor to a phosphate acceptor amino acid residue,  
CC preferably the hydroxyl side chain of a serine or threonine.  
CC Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase  
CC activity is similar to that of a MAPKK such as Raf. has a range of  
CC other activities including a cell growth-regulatory activity, a cell  
CC survival promoting activity, a HAX-1 binding activity, an apoptosis  
CC suppressing activity a MAPKK activation or stimulatory activity, a  
CC nuclear targeting activity and a SRK-specific immunogenic activity.  
CC SRK is useful for identifying agents which modulate cellular  
CC transformations mediated by Ras and SRK and agents that modulate the  
CC apoptosis suppression activity of SRK. This information may be useful  
CC in the treatment of autoimmune diseases, tumours and apoptosis  
CC related disorders.  
XX  
SQ Sequence 14 AA;  
QY Query Match 3.1%; Score 14; DB 21; Length 14;  
Best Local Similarity 100.0%; Pred. No. 7e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 406 AKONSSKTTSKRRG 419  
1 AKONSSKTTSKRRG 14  
RESULT 14  
AAB71969  
ID AAB71969 standard; Peptide; 13 AA.  
XX  
AC AAB71969;  
XX  
DT 11-MAY-2001 (first entry)  
XX  
DE Human TGF-beta receptor serine/threonine kinase domain #2.  
XX





```
XX PA (CIBA ) CIBA GEIGY AG.
XX XX
XX PI Helm J, Seebach P, Takabayashi K;
XX DR WPI; 1992-026641/04.
XX DR N-PSDB; AAQ0547.
XX PT In vitro processing of fusion protein - using yeast endo:protease
XX PT ysc F and carboxypeptidase ysc alpha, for prodn. of biologically
XX PT active mature protein
XX PS Example 8; Page 23; 35pp; English.
XX CC The sequence (SEQ ID NO: 2) is encoded by an insert from plasmid
XX CC pDP34/GAPFL-KEX1 and is a sol. form of ysc alpha. The KEX1
XX CC gene encoding the complete protein (A. Dmochowska et al., Cell 50,
XX CC 573-584 (1987)) was digested with EcoRI and XhoI to isolate a
XX CC fragment encoding a C-terminal truncated sol. protein. The gene
XX CC was expressed in S. cerevisiae. The resulting recombinant sol.
XX CC enzyme can be used, together with similarly produced sol. yscF
XX CC for the prepn. of mature proteins by in vitro processing of
XX CC suitably tailored fusion proteins (see AAR20471, 72).
XX CC See also AAR20469-72.
XX CC
XX SQ Sequence 598 AA;
XX
XX Query Match 2.2%; Score 10; DB 13; Length 598;
XX Best Local Similarity 100.0%; Pred. No. 1.1;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 434 SEGDDDDDD 443
XX |||||
XX Db 564 SEGDDDDDD 573
XX
XX RESULT 17
XX AAR05457
XX ID AAR05457 standard; protein; 729 AA.
XX AC AAR05457;
XX XX
XX DT 17-OCT-1990 (first entry)
XX XX
XX DE KEX1 amino acid sequence.
XX XX
XX KW KEX1; carboxypeptidase; protease; insulin; ds.
XX XX
XX S Saccharomyces cerevisiae.
XX XX
XX PN US4929553-A.
XX XX
XX PD 29-MAY-1990.
XX XX
XX PF 29-MAY-1987; 87US-0055555.
XX XX
XX PR 29-MAY-1987; 87US-0055555.
XX XX
XX PA (CANA ) CANADIAN PATENE & DEV LTD.
XX XX
XX PI Bussey H, Dmochowska A, Thomas DY, Dignard D;
XX XX
XX DR WPI; 1990-192933/25.
XX DR N-PSDB; AAQ04768.
XX XX
XX PT Specific processing of secreted proteins in transformed yeast -
XX PT using vector contg. yeast KEX1 gene, encoding protease, and DNA
XX PT sequence encoding precursor polypeptide eg. Preproinsulin
XX PS Disclosure; ; P; English.
XX CC Gene product is a protease with carboxypeptidase B activity, useful
XX CC in transforming yeast, producing gene products such as prepro
```

```
CC insulin and prepro opio, melanocortin.
XX SQ Sequence 729 AA;
XX
XX Query Match 2.2%; Score 10; DB 11; Length 729;
XX Best Local Similarity 100.0%; Pred. No. 1.3;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 434 SEGDDDDDD 443
XX |||||
XX Db 564 SEGDDDDDD 573
XX
XX RESULT 18
XX AAW01897
XX ID AAW01897 standard; Protein; 764 AA.
XX XX
XX AC AAW01897;
XX XX
XX DT 24-NOV-1996 (first entry)
XX XX
XX DE Nonsense-mediated mRNA decay 2 C-terminal.
XX XX
XX KW Nonsense-mediated mRNA decay; NMD2; Upflp; inhibition.
XX XX
XX OS Saccharomyces cerevisiae.
XX XX
XX PN WO9622301-A1.
XX XX
XX PD 25-JUL-1996.
XX XX
XX PF 27-DEC-1995; 95WO-US16930.
XX XX
XX PR 20-JAN-1995; 95US-0375300.
XX XX
XX PA (UYMA-) UNIV MASSACHUSETTS MEDICAL CENT.
XX XX
XX PI He F, Jacobson AS;
XX XX
XX DR WPI; 1996-354469/35.
XX DR N-PSDB; AAT31994.
XX XX
XX PT Isolated nonsense-mediated mRNA decay gene and protein - used to
XX PT develop methods for inhibiting the decay pathway for producing
XX PT heterologous or endogenous proteins
XX XX
XX PS Claim 10; Page 54-56; 73pp; English.
XX XX
XX CC The NMD2 gene (AAT31993) is named after its role in the Nonsense-
XX CC mediated mRNA Decay pathway. The protein, Nmd2p, binds to Upflp.
XX CC A C-terminal fragment of the protein (AAT31994) also binds Upflp and,
XX CC when overexpressed in the host cell, the fragment inhibits the
XX CC function of Upflp, thereby inhibiting the nonsense-mediated
XX CC mRNA decay pathway.
XX CC
XX SQ Sequence 764 AA;
XX
XX Query Match 2.2%; Score 10; DB 17; Length 764;
XX Best Local Similarity 100.0%; Pred. No. 1.4;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 437 DDDDDDDGEE 446
XX |||||
XX Db 577 DDDDDDDGEE 586
XX
XX RESULT 19
XX AAY98056
XX ID AAY98056 standard; Protein; 764 AA.
XX XX
XX AC AAY98056;
XX XX
XX DT 07-SEP-2000 (first entry)
```

XX DE Yeast NMD2 carboxy terminal fragment.

XX KM Yeast; nonsense-mediated mRNA decay pathway; NMD2; UPF3; UPF1;  
KM upf1p; gene therapy; beta-thalassemia; cancer;  
KM polycystic kidney disease; Duchenne muscular dystrophy; Fanconi anaemia;  
KM haemophilia; hypercholesterolaemia; neurofibromatosis; Tay-Sachs disease;  
KM glycogen storage disease; cystic fibrosis; adenomatous polyposis coli;  
KM Cowden disease; Maple syrup urine disease; Wilson disease;  
KM Niemann-Pick disease; Turcot syndrome; McArdle disease;  
KM ornithine transcarbamylase deficiency; antisense therapy; ribozyme.

XX OS Saccharomyces cerevisiae.

XX PN US6071700-A.

XX PD 06-JUN-2000.

XX PF 21-OCT-1998; 98US-0177431.

XX PR 20-JAN-1995; 95US-0375300.

XX PR 21-OCT-1997; 97US-0955472.

XX PA (UYMA-) UNIV MASSACHUSETTS.

XX PI Jacobson AS, He F;

XX DR WPI; 2000-422078/36.

XX DR N-PSDB; AAA39451.

XX PT Testing for compounds able to modulate the nonsense mediated mRNA decay  
PT pathway for increasing endogenous protein production by incubating a  
PT candidate compound with a cell having a gene with a nonsense mutation  
PT and a gene without mutation -

XX PS Disclosure; Fig 1; 57pp; English.

XX CC The nonsense-mediated mRNA decay pathway has an important role in the  
CC degradation of mRNA transcripts that contain a stop codon. The products  
CC of the UPF1 and UPF3 genes (proteins Upf1p and Upf3p, respectively) are  
CC essential components of this pathway. Mutations in these genes stabilise  
CC mRNA containing premature stop codons. The present sequence is the  
CC carboxy terminal of the NMD2 gene from Saccharomyces cerevisiae. When  
CC the present sequence is overexpressed, it binds to Upf1p to inhibit its  
CC function. Inhibition of the nonsense-mediated mRNA decay pathway is a  
CC useful means of treating disorders caused by the presence of nonsense  
CC mutations, e.g. breast cancer, polycystic kidney diseases, Niemann-Pick  
CC disease, adenomatous polyposis coli, cystic fibrosis, Fanconi's anaemia,  
CC haemophilia, hypercholesterolaemia, neurofibromatosis, ornithine  
CC transcarbamylase deficiency, retinoblastoma, glycogen storage disease,  
CC McArdle disease, colorectal cancer, Tay-Sachs disease, Cowden disease,  
CC Wilson disease, polycystic kidney disease, Duchenne muscular dystrophy,  
CC adrenocortical carcinoma, Li-Fraumeni syndrome, Maple syrup urine  
CC disease, Turcot syndrome or beta-thalassemia. Antisense mRNA molecules  
CC complementary to the NMD2 gene (AAA39450) may be used for antisense  
CC therapy of the above mentioned disorders since the antisense molecules  
CC would inhibit translation of the NMD2 mRNA. Additionally, the NMD2 gene  
CC may be manipulated for ribozyme therapy of the above mentioned  
CC disorders: ribozymes designed to catalytically cleave nonsense-mediated  
CC mRNA decay pathway mRNAs e.g. NMD2 mRNA, can also be used to prevent  
CC translation and therefore expression of these mRNAs.

XX SQ Sequence 764 AA;

Query Match 2.2%; Score 10; DB 21; Length 764;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGEE 446  
DB 577 DDDDDDDGEE 586

RESULT 20

AAW01896 standard; Protein; 1089 AA.

XX AC AAW01896;

XX DT 24-NOV-1996 (first entry)

XX DE Nonsense-mediated mRNA decay 2 protein.

XX KM Nonsense-mediated mRNA decay; NMD2; Upf1p; inhibition.

XX OS Saccharomyces cerevisiae.

XX PN W09622301-A1.

XX PD 25-JUL-1996.

XX PF 27-DEC-1995; 95WO-US16930.

XX PR 20-JAN-1995; 95US-0375300.

XX PA (UYMA-) UNIV MASSACHUSETTS MEDICAL CENT.

XX PI He F, Jacobson AS;

XX DR WPI; 1996-354469/35.

XX DR N-PSDB; AAT31993.

XX PT Isolated nonsense-mediated mRNA decay gene and protein - used to  
PT develop methods for inhibiting the decay pathway for producing  
PT heterologous or endogenous proteins

XX PS Claim 8; Page 49-52; 73pp; English.

XX CC The NMD2 gene (AAT31993) is named after its role in the Nonsense-  
CC Mediated mRNA Decay pathway. The protein, Nmd2p, binds to Upf1p.  
CC A C-terminal fragment of the protein (AAT31994) also binds Upf1p and,  
CC when overexpressed in the host cell, the fragment inhibits the  
CC function of Upf1p, thereby inhibiting the nonsense-mediated  
CC mRNA decay pathway.

XX SQ Sequence 1089 AA;

Query Match 2.2%; Score 10; DB 17; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGEE 446  
DB 902 DDDDDDDGEE 911

RESULT 21

AAW05835 standard; Protein; 1089 AA.

XX AC AAY05835;

XX DT 02-AUG-1999 (first entry)

XX DE Yeast Nmd2p involved in nonsense-mediated mRNA decay pathway.

XX KM NMD2 gene; Nmd2p protein; nonsense-mediated mRNA decay pathway;  
KM Yeast; nonsense mutation; therapy; breast cancer; haemophilia;  
KM polycystic kidney disease; Niemann-Pick disease; cystic fibrosis;  
KM adenomatous polyposis coli; Fanconi's anaemia; neurofibromatosis;  
KM hypercholesterolaemia; ornithine transcarbamylase deficiency;  
KM retinoblastoma; glycogen storage disease; McArdle disease; cancer;  
KM Tay-Sachs disease; Cowden disease; Wilson disease;  
KM beta-thalassemia.

XX OS Saccharomyces cerevisiae.

XX Key Location/Qualifiers  
FH Peptide 26..29  
FT /note= "nuclear localisation signal"  
FT Peptide 42..46  
FT /note= "nuclear localisation signal"  
FT Region 843..975  
FT /note= "acidic region"  
FT Protein 326..1089  
FT /note= "C-terminal fragment, inhibits  
nonsense-mediated mRNA decay"  
XX  
XX WO9920797-A1.  
XX  
XX 29-APR-1999.  
XX  
XX 21-OCT-1998; 98WO-US22365.  
XX  
XX 21-OCT-1997; 97US-0955472.  
XX  
XX (UYMA-) UNIV MASSACHUSETTS.  
XX  
XX He F, Jacobson AS;  
XX  
XX WPI; 1999-302756/25.  
XX  
XX N-PSDB; AAX25601.  
XX  
XX mRNA decay pathway genes and protein useful for treating breast  
XX cancer and cystic fibrosis  
XX  
XX Example 3; Fig 1A-C; 116pp; English.  
XX  
XX This sequence represents Nmd2p, the protein encoded by the newly  
XX discovered NMD2 gene (see AAX25601) of Saccharomyces cerevisiae.  
XX NMD2 was named after its role in the nonsense-mediated mRNA decay  
XX pathway. The Nmd2p protein binds to Upf1, another protein in the  
XX decay pathway. A C-terminal fragment of Nmd2p also binds to Upf1  
XX and, when overexpressed in a host cell, inhibits the function of  
XX Upf1, thereby inhibiting the nonsense-mediated mRNA decay pathway.  
XX The invention also relates to methods of inhibiting the  
XX nonsense-mediated mRNA decay pathway to stabilise mRNA transcripts  
XX containing a nonsense codon which normally would cause an increased  
XX in the transcript decay rate. Such stabilisation allows increased  
XX translation from the transcript, and is therefore useful for the  
XX production of a recombinant protein. The invention also relates to  
XX methods of identifying molecules that inhibit the nonsense-mediated  
XX mRNA decay pathway, and the use of such molecules (e.g. the  
XX C-terminal fragment of Nmd2p or an antisense molecule) for  
XX treatment of disorders associated with nonsense mutations, such  
XX as breast cancer, polycystic kidney disease I and II, Niemann-Pick  
XX disease, adenomatous polyposis coli, cystic fibrosis, Fanconi's  
XX anaemia, haemophilia, hypercholesterolaemia, neurofibromatosis,  
XX ornithine transcarbamylase deficiency, retinoblastoma, glycogen  
XX storage disease, McArdle disease, cancer, Tay-Sachs disease, Cowden  
XX disease, Wilson disease or beta-thalassemia (claimed).  
XX  
SQ Sequence 1089 AA;  
Query Match 2.2%; Score 10; DB 20; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGEE 446  
Db 902 DDDDDDDGEE 911

RESULT 22  
ID AAY98055 standard; Protein; 1089 AA.  
XX AAY98055;  
AC AAY98055;  
XX

DT 07-SEP-2000 (first entry)  
XX  
XX Yeast Nmd2p.  
DE  
XX  
XX Yeast; nonsense-mediated mRNA decay pathway; NMD2; UPF3; UPF1;  
KW upf1p; upf3p; gene therapy; beta-thalassemia; cancer;  
KW polycystic kidney disease; Duchenne muscular dystrophy; Fanconi anaemia;  
KW haemophilia; hypercholesterolaemia; neurofibromatosis; Tay-Sachs disease;  
KW glycogen storage disease; cystic fibrosis; adenomatous polyposis coli;  
KW Cowden disease; Maple syrup urine disease; Wilson disease;  
KW Niemann-Pick disease; Turcot syndrome; McArdle disease;  
KW ornithine transcarbamylase deficiency; antisense therapy; ribozyme.  
XX  
XX Saccharomyces cerevisiae.  
XX  
XX  
XX Key Location/Qualifiers  
FH Peptide 326..1089  
FT /label= Carboxy terminal  
FT /note= "Fragment is described in AAY98056"  
XX  
XX US6071700-A.  
XX  
XX 06-JUN-2000.  
XX  
XX 21-OCT-1998; 98US-0177431.  
XX  
XX 20-JAN-1995; 95US-0375300.  
XX 21-OCT-1997; 97US-0955472.  
XX  
XX (UYMA-) UNIV MASSACHUSETTS.  
XX  
XX Jacobson AS, He F;  
XX  
XX WPI; 2000-422078/36.  
XX  
XX N-PSDB; AAA39450.  
XX  
XX Testing for compounds able to modulate the nonsense mediated mRNA decay  
XX pathway for increasing endogenous protein production by incubating a  
XX candidate compound with a cell having a gene with a nonsense mutation  
XX and a gene without mutation  
XX  
XX Disclosure; Fig 1; 57pp; English.  
XX  
XX The nonsense-mediated mRNA decay pathway has an important role in the  
XX degradation of mRNA transcripts that contain a stop codon. The products  
XX of the UPF1 and UPF3 genes (proteins Upf1 and Upf3p, respectively) are  
XX essential components of this pathway. Mutations in these genes stabilise  
XX mRNA containing premature stop codons. The present sequence is the NMD2  
XX gene protein from Saccharomyces cerevisiae: Nmd2p. The carboxy terminal  
XX of this protein binds to Upf1 to inhibit its function. Inhibition of the  
XX nonsense-mediated mRNA decay pathway is a useful means of treating  
XX disorders caused by the presence of nonsense mutations, e.g. breast  
XX cancer, polycystic kidney diseases, Niemann-Pick disease, adenomatous  
XX polyposis coli, cystic fibrosis, Fanconi's anaemia, haemophilia,  
XX hypercholesterolaemia, neurofibromatosis, ornithine transcarbamylase  
XX deficiency, retinoblastoma, glycogen storage disease, McArdle disease,  
XX colorectal cancer, Tay-Sachs disease, Cowden disease, Wilson disease,  
XX polycystic kidney disease, Duchenne muscular dystrophy, adrenocortical  
XX carcinoma, Li-Fraumeni syndrome, Maple syrup urine disease, Turcot  
XX syndrome or beta-thalassemia. Antisense mRNA molecules complementary to  
XX the NMD2 gene may be used for antisense therapy of the above mentioned  
XX disorders since the antisense molecules would inhibit translation of the  
XX NMD2 mRNA. Additionally, the NMD2 gene may be manipulated for ribozyme  
XX therapy of the above mentioned disorders: ribozymes designed to  
XX catalytically cleave nonsense-mediated mRNA decay pathway mRNAs e.g. NMD2  
XX mRNA, can also be used to prevent translation and therefore expression of  
XX these mRNAs.  
XX  
SQ Sequence 1089 AA;  
Query Match 2.2%; Score 10; DB 21; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



OY 437 DDDDDDDGEE 446  
Db 902 DDDDDDDGEE 911

RESULT 23  
AAG21315  
ID AAG21315 standard; Protein; 133 AA.

XX  
AC AAG21315;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 23829.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.

XX Arabidopsis thaliana.

PN EP1033405-A2.

XX PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 25-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 9; DB 21; Length 133;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches \* 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 SEGDDDDDD 442  
Db 47 SEGDDDDDD 55

RESULT 24

AAG21314  
ID AAG21314 standard; Protein; 156 AA.  
XX AC AAG21314;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 23828.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 9; DB 21; Length 156;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 SEGDDDDDD 442  
Db 70 SEGDDDDDD 78

RESULT 25  
AAG07429 standard; Protein; 201 AA.

XX AC AAG07429;  
XX DT 17-OCT-2000 (first entry)  
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 4579.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146389.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149175.  
PR 20-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 25-AUG-1999; 99US-0149930.  
PR 26-AUG-1999; 99US-0150566.  
PR 27-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.



PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 9; DB 21; Length 201;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 434 SEGDDDDDD 442  
Db 47 SEGDDDDDD 55

RESULT 26.

AAAG07428  
ID AAG07428 standard; Protein; 224 AA.

XX AC AAG07428;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 4578.

XX KW protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.

OS Arabidopsis thaliana.

XX PN EPI033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 27-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.

PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 26-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 9; DB 21; Length 224;  
Best Local Similarity 100.0%; Pred.No. 4.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 434 SEGDDDDDD 442  
DB 70 SEGDDDDDD 78

RESULT 27  
AAR55371  
ID AAR55371 standard; Protein; 502 AA.  
XX  
AC AAR55371;  
XX  
DT 20-JAN-1995 (first entry)  
XX  
DE Mouse Activin receptor-like kinase 1 (MALK-1).  
XX  
KW serine threonine kinases; activin receptors; Act-R; superfamily;  
KW transforming growth factor; TGF; diagnostics; detection; therapy;  
KW rheumatoid arthritis; glomerular nephritis; fibrosis.  
XX  
OS Mus musculus.  
XX  
PN WO9411502-A.  
XX  
PD 26-MAY-1994.  
XX  
PF 17-NOV-1993; 93WO-GB02367.  
XX

PR 17-NOV-1992; 92GB-0024057.  
PR 08-MAR-1993; 93GB-0004677.  
PR 08-MAR-1993; 93GB-0004680.  
PR 28-MAY-1993; 93GB-0011047.  
PR 02-JUL-1993; 93GB-0013763.  
PR 03-AUG-1993; 93GB-0016099.  
PR 15-OCT-1993; 93GB-0021344.  
XX  
PA (LUDW-) \*LUDWIG INST CANCER RES.  
XX  
PI Dijke P, Franzen P, Heldin C, Miyazono K, Yamashita H;  
XX  
DR WPI; 1994-183503/22.  
DR N-PSDB; AAQ66671.  
XX  
PT Activin receptor-like kinase(s) with serine/threonine kinase  
PT domains - have activin/TGF beta-type I receptor function and can  
PT be used in diagnosis or therapy of rheumatoid arthritis,  
PT glomerular nephritis, fibrosis, etc.  
XX  
XX Claim 3; Page 60-61; 97pp; English.  
XX  
CC The inventors have identified a new family of receptor kinases  
CC called activin receptor-like kinases (ALK). Their discovery was  
CC based on the realisation that receptor serine/threonine kinases  
CC form a new receptor family, which may include the type II receptors  
CC for other proteins in the transforming growth factor (TGF) beta  
CC superfamily. The activin receptor type II sequences from mouse and  
CC the daf1 gene product of C. elegans have high sequence similarity  
CC and were used to design degenerate primers to clone related CDNA's  
CC (see AAQ66643-49). Six distinct putative receptor serine/threonine  
CC kinases were identified, called ALK (human ALK proteins are shown  
CC in AAR55366-70, mouse ALK are shown in AAR55371-74). Products of the  
CC invention can be used in therapy, eg. to modulate conditions  
CC associated with activin or TGF beta activity. These conditions  
CC include fibrosis, eg. liver cirrhosis and pulmonary fibrosis, cancer,  
CC rheumatoid arthritis and glomeronephritis.  
XX  
SQ Sequence 502 AA;  
XX  
Query Match 2.0%; Score 9; DB 15; Length 502;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 131 HRDLKSRNV 139  
Db 327 HRDLKSRNV 335  
XX  
RESULT 28  
AY33304  
ID AAY33304 standard; Protein; 502 AA.  
XX  
AC AAY33304;  
XX  
DT 26-NOV-1999 (first entry)  
XX  
DE Human mALK-1 clone AM6 protein.  
XX  
KW mALK-1; activin receptor-like kinase; therapy; activin; TGF-beta; cancer;  
KW fibrosis; liver cirrhosis; pulmonary fibrosis; glomerulonephritis; mouse;  
KW rheumatoid arthritis; detection; diagnosis; drug screening.  
XX  
OS Mus sp.  
XX  
PN WO9946386-A1.  
XX  
PD 16-SEP-1999.  
XX  
PF 12-MAR-1999; 99WO-US05735.  
XX  
PR 13-MAR-1998; 98US-0039177.  
XX

PA (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI Miyazono K, Imamura T, Ten Dijke P;  
XX  
DR WPI; 1999-551413/46.  
DR N-PSDB; AAZ09846.  
XX  
PT New isolated activin receptor-like kinases, used to develop products  
PT for treating e.g. fibrosis, cancer, rheumatoid arthritis and  
PT glomerulonephritis  
XX  
PS Disclosure; Page 86-88; 110pp; English.  
XX  
CC This invention describes novel human and murine activin receptor-like  
CC kinases (ALK's). The novel ALK products can be used in therapy, e.g. to  
CC modulate conditions associated with activin or TGF-beta activity, such as  
CC fibrosis, e.g. liver cirrhosis and pulmonary fibrosis, cancer, rheumatoid  
CC arthritis and glomerulonephritis. The products can also be used for  
CC detection, diagnosis and drug screening. This sequence represents the  
CC murine mALK-1 protein isolated from clone AM6.  
XX  
SQ Sequence 502 AA;  
XX  
Query Match 2.0%; Score 9; DB 20; Length 502;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 131 HRDLKSRNV 139  
Db 327 HRDLKSRNV 335  
XX  
RESULT 29  
AAR41922  
ID AAR41922 standard; Protein; 505 AA.  
XX  
AC AAR41922;  
XX  
DT 20-APR-1994 (first entry)  
XX  
DE MISR3.  
XX  
KW MIS, Mullerian Inhibitory Substance; receptor;  
KW transforming growth factor; inhibin; BMP;  
KW membrane serine/threonine kinase receptor;  
KW bone morphogenesis protein.  
XX  
OS Rattus rattus.  
XX  
PN WO9319177-A.  
XX  
PD 30-SEP-1993.  
XX  
PF 15-MAR-1993; 93WO-US02387.  
XX  
PR 18-MAR-1992; 92US-0853396.  
PR 11-MAR-1993; 93US-0029673.  
XX  
PA (GENO ) GEN HOSPITAL CORP.  
XX  
PI Donahoe PK, Gustafson M, He WW;  
XX  
DR WPI; 1993-320743/40.  
DR N-PSDB; AAQ49765.  
XX  
PT New receptors of the transforming growth factor-beta receptor  
PT family - comprising Mullerian Inhibitory Substance Receptors and  
PT inhibin receptors  
XX  
PS Claim 51; Fig 3; 59pp; English.  
XX  
CC Misr1 (AAQ49763) is believed to encode an isoform of the rat  
CC MIS receptor. Misr2A/misr2B (AAQ49764), misr3 (AAQ49765) and misr4





CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins.

XX Sequence 648 AA;

Query Match 2.0%; Score 9; DB 23; Length 648;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 51 EKEAEILSV 59  
Db 303 EKEAEILSV 311

RESULT 32

ABP27599 ID ABP27599 standard; Protein; 652 AA.

XX ABP27599;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 4374.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
KM antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus agalactiae.

PN WO200234771-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB04789.

PR 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

PR 07-MAR-2001; 2001GB-0005640.

PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;

PI Tettelin H;

XX MPI; 2002-352536/38.

XX N-PSDB; ABN68230.

PS Claim 1; Page 3598; 4525bp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins.

XX Sequence 652 AA;

XX Query Match 2.0%; Score 9; DB 23; Length 652;

Best Local Similarity 100.0%; Pred. No. 11;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 51 EKEAEILSV 59  
Db 307 EKEAEILSV 315

RESULT 33

ABP27600 ID ABP27600 standard; Protein; 652 AA.

XX ABP27600;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 4376.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
KM antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus pyogenes.

PN WO200234771-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB04789.

PR 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

PR 07-MAR-2001; 2001GB-0005640.

PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;

PI Tettelin H;

XX MPI; 2002-352536/38.

XX N-PSDB; ABN68231.

PS Claim 1; Page 3598; 4525bp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins.

```
Query Match      2.0%; Score 9; DB 23; Length 652;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 51 EKEAEILSV 59
Db 307 EKEAEILSV 315

RESULT 34
AAE16259
ID AAE16259 standard; Protein; 656 AA.
XX
AC AAE16259;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human kinase PKIN-5 protein.

Human; kinase; PKIN-5; cancer; leukaemia; adenocarcinoma; osteoarthritis;
immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;
Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;
allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;
autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoporosis;
Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;
rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;
hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;
cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;
congestive heart failure; ischaemic heart disease; lung tumour; gout;
fatty liver; Niemann-Pick's disease; gene therapy.

XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers.
FT Domain 13..252
FT /label= Protein_kinase_domain
FT Domain 14..197
FT /label= Protein_kinase_domain
FT Domain 14..257
FT /note= "Eukaryotic protein kinase domain"
FT Domain 14..252
FT /label= Protein_kinase_domain
FT Domain 14..253
FT /label= Protein_kinase_domain
XX
PN WO200196547-A2.
XX
20-DEC-2001.
XX
PF 14-JUN-2001; 2001WO-US19444.
XX
PR 15-JUN-2000; 2000US-212073P.
PR 23-JUN-2000; 2000US-213467P.
PR 30-JUN-2000; 2000US-215651P.
PR 07-JUL-2000; 2000US-21605P.
PR 13-JUL-2000; 2000US-218372P.
PR 25-AUG-2000; 2000US-228056P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;
PI Gandhi AR, Tribouley CM, Walla NK, Yao MG, Lu DAM, Greenwald SR;
PI Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YT;
PI Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;
PI Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JL, Ding L;
PI Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;
XX
DR WPI; 2002-090207/12.
DR N-PSDB; AAD26452.
XX
PT New polypeptides, useful for diagnosing, treating or preventing
PT disorders of growth and development, cardiovascular and lipid, and
```

```
PT diseases such as cancer, comprise human kinase polypeptides -
XX
XX Claim 1; Page 143-144; 197pp; English.
PS
XX
CC The invention relates to human kinase PKIN proteins and their
CC corresponding cDNAs. A composition containing PKIN agonist is useful for
CC treating a disease or condition associated with decreased expression of
CC PKIN and a composition comprising PKIN antagonist is useful for treating
CC a disease or condition associated with overexpression of PKIN. The
CC disorders include cancer (leukaemia, adenocarcinoma, lymphoma, melanoma,
CC myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder
CC (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,
CC atherosclerosis, anaemia, allergies, adult respiratory distress syndrome,
CC autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes
CC mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,
CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,
CC rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,
CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)
CC growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,
CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio
CC vascular disease (arteriovenous fistula, hypertension, vasculitis,
CC aneurysms, congestive heart failure, angina pectoris, myocarditis,
CC ischaemic heart disease, chronic bronchitis, lung tumours); lipid
CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,
CC hypcholesterolaemia, obesity). PKIN DNA is useful for assessing
CC toxicity of a test compound and in gene therapy. The present sequence
CC is human PKIN-5 protein.
XX
SQ Sequence 656 AA;

Query Match      2.0%; Score 9; DB 23; Length 656;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 GSFGSVYRA 33
Db 17 GSFGSVYRA 25

RESULT 35
AAU03554
ID AAU03554 standard; Protein; 909 AA.
XX
AC AAU03554;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human protein kinase #54.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32085.
XX
PR 24-NOV-1999; 99US-0167482.
XX
PA (SUGB-) SUGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX
DR WPI; 2001-343950/36.
DR N-PSDB; AAS06754.
XX
PT Nucleic acids encoding human kinase polypeptides, useful for preventing
```

PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
PT neuronal-associated diseases, and microbial infections -  
XX  
PS Claim 7; Figure 2; 433pp; English.  
XX  
CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The  
CC novel protein kinases have been identified as members of the tyrosine  
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides  
CC encoding protein kinases and the polypeptides may be used in the  
CC prevention, diagnosis and treatment of diseases associated with  
CC inappropriate kinase expression. For example, they may be used to treat  
CC cancers (especially cancers of haematopoietic origin), cardiovascular  
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
CC immune related diseases (e.g. rheumatoid arthritis), neurological  
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
CC parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
CC Additionally, polynucleotides encoding protein kinases may be  
CC used for gene therapy and as DNA probes in diagnostic assays.  
CC The protein kinase polypeptides may be used as antigens in the production  
CC of antibodies against the protein kinases and in assays to identify  
CC modulators of protein kinase expression and activity.

XX SQ Sequence 909 AA;

Query Match 2.0%; Score 9; DB 22; Length 909;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 GSFGSVYRA 33  
|||  
Db 403 GSFGSVYRA 411

RESULT 36

AAR55694  
ID AAR55694 standard; Protein; 2391 AA.

XX AC AAR55694;

DT 06-DEC-1994 (first entry)

XX DE Carbamoyl-phosphate-synthetase II.

XX KW Carbamoyl-phosphate-synthetase II; CPSII; pSCPSII gene;  
XX malaria.

XX OS Plasmodium falciparum.

XX Key Location/Qualifiers

FT Domain 1..690 /note= "glutamine-amidotransferase domain"

FT Domain 1..270 /note= "structural subdomain"

FT Peptide 271..482 /note= "insert sequence"

FT Domain 483..690 /note= "glutaminase subdomain"

FT Domain 691..2391 /note= "carbamoyl-phosphate-synthase domain"

FT Domain 691..1254 /note= "ATP binding subdomain CPsa"

FT Peptide 1255..1857 /note= "insert sequence"

FT Domain 1858..2391 /note= "ATP binding subdomain CPsb"

XX PN WO9412643-A.

XX PD 09-JUN-1994.

XX PF 02-DEC-1993; 93WO-AU00617.

XX

PR 03-DEC-1992; 92AU-0006206.  
PR 16-DEC-1992; 92AU-0006380.

XX PA (UNIX ) UNISEARCH LTD.

XX PI Flores MV, Osullivan WJ, Stewart TS;

XX DR WPI; 1994-200271/24.

XX DR N-PSDB; AAQ62924.

XX PT Nucleic acid encoding carbamoyl phosphate synthetase II -  
XX isolated from Plasmodium falciparum, used to develop prods. for  
XX the treatment of malaria.

XX PS Disclosure; Page 6-16; 31pp; English.

XX CC The cDNA sequence encoding the carbamoyl-phosphate-transferase II  
CC (CPSII) of Plasmodium falciparum was determined. The cDNA encodes  
CC a protein that includes 2 insert sequences not found in other CPSII  
CC proteins. The first separates the putative structural subdomain and  
CC the glutaminase subdomain of the glutamine-amidotransferase subunit  
CC of CPSII, while the second separates 2 ATP binding subdomains of the  
CC CPSII subunit, CPsa and CPsb.

XX SQ Sequence 2391 AA;

Query Match 2.0%; Score 9; DB 15; Length 2391;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGE 445  
|||  
Db 1786 DDDDDDDGE 1794

RESULT 37

AAB83387  
ID AAB83387 standard; peptide; 20 AA.

XX AC AAB83387;

DT 26-MAR-2002 (first entry)

XX DE Lunasin fragment #8.

XX KW Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
XX prostate; liver; kidney; therapy; soybean.

XX OS Glycine max.

XX PN WO200172784-A2.

XX PD 04-OCT-2001.

XX PF 23-MAR-2001; 2001WO-US09453.

XX PR 24-MAR-2000; 2000US-0534705.

XX PA (FILG-) FILGEN BIOSCIENCES INC.

XX PI Galvez AF;

XX DR WPI; 2001-648438/74.

XX PT Treatment or prevention of cancer by administering a lunasin peptide  
XX which binds to non-acetylated histones H3 and H4 and prevents their  
XX acetylation in hypoacetylated repressed chromatin -

XX PS Claim 7; Page 42; 49pp; English.

XX CC This sequence represents a lunasin peptide used in the method of the  
XX invention. The method is for the treatment or prevention of cancer,  
XX comprising administration of a lunasin peptide or its active fragment or

CC analog. The invention is used to treat or prevent cancer, particularly  
CC in a human. The cancers that can be treated and prevented include  
CC those of the colon, upper gastrointestinal tract, breast, prostate,  
CC liver, kidney or any other internal organs or tissues.  
XX  
SQ Sequence 20 AA;

Query Match 1.8%; Score 8; DB 22; Length 20;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDDDD 443  
|||  
Db 13 GDDDDDDDD 20

RESULT 38  
AAB62621  
ID AAB62621 standard; peptide; 20 AA.

AAB62621;

DT 23-JUL-2001 (first entry)

DE Soybean lunasin peptide variant (residues 22-41).

KW Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.

OS Glycine max.

PN WO200134808-A2.

PD 17-MAY-2001.

PF 12-NOV-2000; 2000WO-US31211.

PR 12-NOV-1999; 99US-0165334.

PA (FILG-) FILGEN BIOSCIENCES INC.

PI Galvez AF;

DR WPI; 2001-343605/36.

PT Large scale production of lunasin, a cancer preventive peptide from  
PT soybean, by preparing lunasin gene constructs using Pichia expression  
PT vectors, optimizing lunasin expression parameters and purifying lunasin

Claim 7; Page 62; 67pp; English.

CC The invention provides a method of producing lunasin, a cancer preventive  
CC peptide from soybean, or its active variant by recombinant DNA technology  
CC in large quantities. The method comprises preparing lunasin gene  
CC constructs or its variants using protein expression vectors, optimizing  
CC lunasin expression parameters for large-scale production and performing  
CC a series of isolation and purification to obtain large quantities of  
CC biologically active recombinant lunasin peptide. The lunasin peptide  
CC exhibits inhibitory effects against malignant transformation of cells  
CC induced by chemical carcinogens and viral oncogenes. Sequences  
CC AAB62614-24 represent soybean lunasin variants.

SQ Sequence 20 AA;

Query Match 1.8%; Score 8; DB 22; Length 20;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDDDD 443  
|||  
Db 13 GDDDDDDDD 20

RESULT 39

AAB83386

ID AAB83386 standard; peptide; 21 AA.

AC AAB83386;

DT 26-MAR-2002 (first entry)

DE Lunasin fragment #7.

KW Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
KW prostate; liver; kidney; therapy; soybean.

OS Glycine max.

PN WO200172784-A2.

PD 04-OCT-2001.

PF 23-MAR-2001; 2001WO-US09453.

PR 24-MAR-2000; 2000US-0534705.

PA (FILG-) FILGEN BIOSCIENCES INC.

PI Galvez AF;

DR WPI; 2001-648438/74.

PT Treatment or prevention of cancer by administering a lunasin peptide  
PT which binds to non-acetylated histones H3 and H4 and prevents their  
PT acetylation in hypoacetylated repressed chromatin

PS Claim 7; Page 42; 49pp; English.

CC This sequence represents a lunasin peptide used in the method of the  
CC invention. The method is for the treatment or prevention of cancer,  
CC comprising administration of a lunasin peptide or its active fragment or  
CC analog. The invention is used to treat or prevent cancer, particularly  
CC in a human. The cancers that can be treated and prevented include  
CC those of the colon, upper gastrointestinal tract, breast, prostate,  
CC liver, kidney or any other internal organs or tissues.

SQ Sequence 21 AA;

Query Match 1.8%; Score 8; DB 22; Length 21;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDDDD 443  
|||  
Db 13 GDDDDDDDD 20

RESULT 40

AAB83393

ID AAB83393 standard; peptide; 21 AA.

AC AAB83393;

DT 26-MAR-2002 (first entry)

DE Lunasin related peptide #2.

KW Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
KW prostate; liver; kidney; therapy; soybean.

OS Glycine max.

PN WO200172784-A2.

PD 04-OCT-2001.

XX



PF 23-MAR-2001; 2001WO-US09453.  
XX  
XX 24-MAR-2000; 2000US-0534705.  
PR  
XX (FILG-) FILGEN BIOSCIENCES INC.  
XX  
XX Galvez AF;  
PI  
XX WPI; 2001-648438/74.  
DR  
XX Treatment or prevention of cancer by administering a lunasin peptide  
PT which binds to non-acetylated histones H3 and H4 and prevents their  
PT acetylation in hypoacetylated repressed chromatin  
PS  
XX Disclosure; Page 46; 49pp; English.  
XX  
XX This sequence represents a lunasin related peptide used in the method of  
CC the invention. The method is for the treatment or prevention of cancer,  
CC comprising administration of a lunasin peptide or its active fragment or  
CC analog. The invention is used to treat or prevent cancer, particularly  
CC in a human. The cancers that can be treated and prevented include  
CC those of the colon, upper gastrointestinal tract, breast, prostate,  
CC liver, kidney or any other internal organs or tissues.  
XX  
XX Sequence 21 AA;  
SQ  
  
Query Match 1.8%; Score 8; DB 22; Length 21;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 436 GDDDDDDDD 443  
DB 12 GDDDDDDDD 19  
  
RESULT 41  
AAB62620  
ID AAB62620 standard; peptide; 21 AA.  
XX  
XX AAB62620;  
AC  
XX 23-JUL-2001 (first entry)  
DT  
XX Soybean lunasin peptide variant (residues 22-42).  
DE  
XX Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.  
XX  
XX Glycine max.  
OS  
XX WO200134808-A2.  
XX  
XX 17-MAY-2001.  
PD  
XX 12-NOV-2000; 2000WO-US31211.  
PF  
XX 12-NOV-1999; 99US-0165334.  
PR  
XX (FILG-) FILGEN BIOSCIENCES INC.  
PA  
XX Galvez AF;  
PI  
XX WPI; 2001-343605/36.  
DR  
XX Large scale production of lunasin, a cancer preventive peptide from  
XX soybean, by preparing lunasin gene constructs using Pichia expression  
PT vectors, optimizing lunasin expression parameters and purifying lunasin  
PT  
PS Claim 7; Page 62; 67pp; English.  
XX  
XX The invention provides a method of producing lunasin, a cancer preventive  
CC peptide from soybean, or its active variant by recombinant DNA technology  
CC in large quantities. The method comprises preparing lunasin gene

CC constructs or its variants using protein expression vectors, optimizing  
CC lunasin expression parameters for large-scale production and performing  
CC a series of isolation and purification to obtain large quantities of  
CC biologically active recombinant lunasin peptide. The lunasin peptide  
CC exhibits inhibitory effects against malignant transformation of cells  
CC induced by chemical carcinogens and viral oncogenes. Sequences  
CC AAB62614-24 represent soybean lunasin variants.  
XX  
XX Sequence 21 AA;  
SQ  
  
Query Match 1.8%; Score 8; DB 22; Length 21;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 436 GDDDDDDDD 443  
DB 13 GDDDDDDDD 20  
  
RESULT 42  
AAB83385  
ID AAB83385 standard; peptide; 22 AA.  
XX  
XX AAB83385;  
AC  
XX 26-MAR-2002 (first entry)  
DT  
XX Lunasin fragment #6.  
DE  
XX Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
XX Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
XX prostate; liver; kidney; therapy; soybean.  
XX  
XX Glycine max.  
OS  
XX WO200172784-A2.  
XX  
XX 04-OCT-2001.  
PD  
XX 23-MAR-2001; 2001WO-US09453.  
PF  
XX 24-MAR-2000; 2000US-0534705.  
PR  
XX (FILG-) FILGEN BIOSCIENCES INC.  
PA  
XX Galvez AF;  
PI  
XX WPI; 2001-648438/74.  
DR  
XX Treatment or prevention of cancer by administering a lunasin peptide  
PT which binds to non-acetylated histones H3 and H4 and prevents their  
PT acetylation in hypoacetylated repressed chromatin  
PS  
XX Claim 7; Page 42; 49pp; English.  
XX  
XX This sequence represents a lunasin peptide used in the method of the  
CC invention. The method is for the treatment or prevention of cancer,  
CC comprising administration of a lunasin peptide or its active fragment or  
CC analog. The invention is used to treat or prevent cancer, particularly  
CC in a human. The cancers that can be treated and prevented include  
CC those of the colon, upper gastrointestinal tract, breast, prostate,  
CC liver, kidney or any other internal organs or tissues.  
XX  
XX Sequence 22 AA;  
SQ  
  
Query Match 1.8%; Score 8; DB 22; Length 22;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 436 GDDDDDDDD 443  
DB 13 GDDDDDDDD 20

RESULT 43  
ABB30607  
ID ABB30607 standard; Peptide; 22 AA.  
XX  
AC ABB30607;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Peptide #3258 encoded by breast cell single exon nucleic acid probe.  
XX  
KW Human; microarray; single exon probe; gene expression; breast;  
KW disease; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157271-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00662.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-496933/54.  
XX  
PT New spatially-addressable set of single exon nucleic acid probes,  
PT useful for measuring gene expression in sample derived from human  
PT breast, comprises number of single exon nucleic acid probes  
XX  
PS Claim 27; SEQ ID NO 13575; 327pp + sequence listing; English.  
XX  
CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BT 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a peptide encoded by a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 22 AA;  
QY 437 DDDDDDDG 444  
Db 7 DDDDDDDG 14

Query Match 1.8%; Score 8; DB 22; Length 22;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 44

ABB35771  
ID ABB35771 standard; Peptide; 22 AA.  
XX  
AC ABB35771;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #3277 encoded by human foetal liver single exon probe.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483447/52.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human foetal liver  
XX  
PS Claim 27; SEQ ID NO 28406; 639pp + sequence listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC foetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 22 AA;  
QY 437 DDDDDDDG 444  
Db 7 DDDDDDDG 14

Query Match 1.8%; Score 8; DB 22; Length 22;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 45  
ABB21195  
ID ABB21195 standard; Protein; 22 AA.  
XX  
AC ABB21195;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Protein #3194 encoded by probe for measuring heart cell gene expression.  
XX  
KW Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX

```

OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts
PS Claim 15; SEQ ID No 22965; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 22 AA;
OY
OY 437 DDDDDDDG 444
   |||||
   7 DDDDDDDG 14
RESULT 46
AAM56580
ID AAM56580 standard; Protein; 22 AA.
AC AAM56580;
XX
DT 05-NOV-2001 (first entry)
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 28685.
XX Human, brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX

```

```

PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX      Penn SG, Hanzel DK, Chen W, Rank DR;
PI
DR      WPI; 2001-483446/52.
XX
PT      Single exon nucleic acid probes for analyzing gene expression in human
PT      brains -
PS      Example 4; SEQ ID NO: 28685; 650bp + Sequence Listing; English.
XX
CC      The present invention provides a number of single exon nucleic acid
CC      probes which are derived from genomic sequences expressed in the human
CC      brain. They can be used to measure gene expression in brain cell samples,
CC      which may enable the diagnosis and improved treatment of nervous system
CC      diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC      epilepsy and cancers. The present sequence is a protein encoded by one of
CC      the probes of the invention.
XX
SQ      Sequence    22 AA;
OY      437 DDDDDDDG 444
DB      7 DDDDDDDG 14
        |||||
Query Match          1.8%; Score 8; DB 22; Length 22;
Best Local Similarity 100.0%; Pred.No. 5;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0.

RESULT 47
AAM68960
ID      AAM68960 standard; Protein; 22 AA.
AC      AAM68960;
XX
DT      06-NOV-2001 (first entry)
DE      Human bone marrow expressed probe encoded protein SEQ ID NO: 29266.
XX
XX      Human, bone marrow expressed exon; gene expression analysis; probe;
KW      microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS      Homo sapiens.
XX
PN      WO200157276-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US00668.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX      Penn SG, Hanzel DK, Chen W, Rank DR;
PI
DR      WPI; 2001-488900/53.
XX
```

PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human bone marrow -  
XX  
PS Example 4; SEQ ID NO: 29266; 658bp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention.  
XX  
SQ Sequence 22 AA;

Query Match 1.8%; Score 8; DB 22; Length 22;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 437 DDDDDDDG 444  
D 7 DDDDDDDG 14

RESULT 48

AAM16787  
ID AAM16787 standard; Protein; 22 AA.  
XX  
AC AAM16787;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE Peptide #3221 encoded by probe for measuring cervical gene expression.  
XX  
KW Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157278-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00670.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488901/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
XX  
PS Claim 27; SEQ ID No 21613; 487bp; English.

CC The present invention relates to human single exon nucleic acid probes  
CC (SENP: see AAI10068-AA128459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human Hela cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 22 AA;

Query Match 1.8%; Score 8; DB 22; Length 22;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 437 DDDDDDDG 444  
Db 7 DDDDDDDG 14

RESULT 49

AAM29271  
ID AAM29271 standard; Protein; 22 AA.  
XX  
AC AAM29271;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Peptide #3308 encoded by probe for measuring placental gene expression.  
XX  
KW Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00663.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488897/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
PS Claim 27; SEQ ID No 29540; 654bp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes (SENP:  
CC see AAI31315-AA157546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.  
XX  
SQ Sequence 22 AA;

Query Match 1.8%; Score 8; DB 22; Length 22;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
Db 7 DDDDDDDG 14

RESULT 50

AAM04504



ID AAM04504 standard; Protein; 22 AA.  
XX  
AC AAM04504;  
XX  
DT 09-OCT-2001 (first entry)  
XX  
DE Peptide #3186 encoded by probe for measuring breast gene expression.  
XX  
KW Probe; human; breast disease; breast cancer; development disorder;  
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200157270-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 29-JAN-2001; 2001WO-US00661.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX  
PR 26-MAY-2000; 2000US-0207456.  
XX  
PR 30-JUN-2000; 2000US-0608408.  
XX  
PR 03-AUG-2000; 2000US-0632366.  
XX  
PR 21-SEP-2000; 2000US-0234687.  
XX  
PR 27-SEP-2000; 2000US-0236359.  
XX  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-476286/51.  
XX  
PT Novel single exon nucleic acid probe used to measuring gene expression  
PT in a human breast -  
XX  
PS Claim 27; SEQ ID NO 13244; 322pp; English.  
XX  
CC The present invention relates to novel single exon nucleic acid probes  
CC (see AAI00010-AAI10067). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for measuring human gene expression in  
CC a human breast sample, where the probe hybridises at high stringency to a  
CC nucleic acid expressed in the human breast. The probes are useful for  
CC predicting, diagnosing, grading, staging, monitoring and prognosing  
CC diseases of the human breast, particularly those diseases with polygenic  
CC aetiology. The diseases include: breast cancer, disorders of development,  
CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
CC breast disease and non-carcinoma tumours.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 22 AA;  
XX  
Query Match 1.8%; Score 8; DB 22; Length 22;  
Best Local Similarity 100.0%; Pred. NO. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDDG 444  
|||  
DB 7 DDDDDDDG 14

KW Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.  
XX  
OS Glycine max.  
XX  
PN WO200134808-A2.  
XX  
PD 17-MAY-2001.  
XX  
PF 12-NOV-2000; 2000WO-US31211.  
XX  
PR 12-NOV-1999; 99US-0165334.  
XX  
PR (FILG-) FILGEN BIOSCIENCES INC.  
XX  
PA Galvez AF;  
XX  
PI WPI; 2001-343605/36.  
XX  
DR Large scale production of lunasin, a cancer preventive peptide from  
XX soybean, by preparing lunasin gene constructs using Pichia expression  
XX vectors, optimizing lunasin expression parameters and purifying lunasin  
XX  
PT Claim 7; Page 62; 67pp; English.  
XX  
PS The invention provides a method of producing lunasin, a cancer preventive  
XX peptide from soybean, or its active variant by recombinant DNA technology  
XX in large quantities. The method comprises preparing lunasin gene  
XX constructs or its variants using protein expression vectors, optimizing  
XX lunasin expression parameters for large-scale production and performing  
XX a series of isolation and purification to obtain large quantities of  
XX biologically active recombinant lunasin peptide. The lunasin peptide  
XX exhibits inhibitory effects against malignant transformation of cells  
XX induced by chemical carcinogens and viral oncogenes. Sequences  
XX AAB62614-24 represent soybean lunasin variants.  
XX  
SQ Sequence 22 AA;  
XX  
Query Match 1.8%; Score 8; DB 22; Length 22;  
Best Local Similarity 100.0%; Pred. NO. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDDDD 443  
|||  
DB 13 GDDDDDDDD 20

RESULT 52  
ABG38546  
ID ABG38546 standard; Peptide; 22 AA.  
XX  
AC ABG38546;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human peptide encoded by genome-derived single exon probe SEQ ID 28211.  
XX  
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;  
KW chronic obstructive pulmonary disease; interstitial lung disease;  
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;  
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
KW primary ciliary dyskinesia; pulmonary hypertension;  
KW hyaline membrane disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200186003-A2.  
XX  
PD 15-NOV-2001.  
XX

PF 30-JAN-2001; 2001WO-US00665.  
XX  
PR 04-FEB-2000; 2000US-180312P.  
PR 26-MAY-2000; 2000US-207456P.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-234687P.  
PR 27-SEP-2000; 2000US-236359P.  
PR 04-OCT-2000; 2000GB-0024263.  
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX

DR WPI; 2002-114183/15.

PT Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples -

Claim 27; SEQ ID NO 28211; 634bp; English.

CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of  
CC probes; the novel set of probes which hybridise at high stringency to a  
CC nucleic acid expressed in the human lung; measuring gene expression in a  
CC sample derived from human lung, comprising (a) contacting the array with  
CC a collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of  
CC the array; identifying exons in a eukaryotic genome, comprising  
CC (a) algorithmically predicting at least one exon from genomic sequences  
CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarray having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene  
CC expression analysis, and for identifying exons in a gene, particularly  
CC using human lung derived mRNA and for the study of lung diseases  
CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic  
CC and hyaline membrane disease. The present sequence is a peptide/protein  
CC encoded by a single exon probe of the invention.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 22 AA;

Query Match 1.8%; Score 8; DB 23; Length 22;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
DB 7 DDDDDDDG 14

RESULT 53

AAB83395

ID AAB83395 standard; peptide; 25 AA.

XX AAB83395;

AC 26-MAR-2002 (first entry)

DT 26-MAR-2002 (first entry)

XX Lunasin related peptide #4.

DE Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
KW prostate; liver; kidney; therapy; soybean.

XX Glycine max.

OS WO200172784-A2.

PN 04-OCT-2001.

PD 23-MAR-2001; 2001WO-US09453.

PF 24-MAR-2000; 2000US-0534705.

PR (FILG-) FILGEN BIOSCIENCES INC.

XX Galvez AF;

PI WPI; 2001-648438/74.

DR Treatment or prevention of cancer by administering a lunasin peptide  
XX which binds to non-acetylated histones H3 and H4 and prevents their  
PT acetylation in hypacetylated repressed chromatin -

PT Disclosure; Page 47; 49pp; English.

XX This sequence represents a lunasin related peptide used in the method of  
CC the invention. The method is for the treatment or prevention of cancer,  
CC comprising administration of a lunasin peptide or its active fragment or  
CC analog. The invention is used to treat or prevent cancer, particularly  
CC in a human. The cancers that can be treated and prevented include  
CC those of the colon, upper gastrointestinal tract, breast, prostate,  
CC liver, kidney or any other internal organs or tissues.

XX  
SQ Sequence 25 AA;

Query Match 1.8%; Score 8; DB 22; Length 25;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDD 443

DB 16 GDDDDDDD 23

RESULT 54

AAB29408

ID AAB29408 standard; Peptide; 31 AA.

XX AAB29408;

AC 01-FEB-2002 (first entry)

DE Peptide #2059 encoded by breast cell single exon nucleic acid probe.

XX Human; microarray; single exon probe; gene expression; breast;  
KW disease; cancer.

XX Homo sapiens.

OS WO200157271-A2.

PN 09-AUG-2001.

PD  
XX

PF 30-JAN-2001; 2001WO-US00662.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-496933/54.  
XX  
PT New spatially-addressable set of single exon nucleic acid probes,  
PT useful for measuring gene expression in sample derived from human  
PT breast, comprises number of single exon nucleic acid probes -  
XX  
XX Claim 27; SEQ ID NO 12376; 327pp + sequence listing; English.  
XX  
CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BT 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a peptide encoded by a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 31 AA;  
XX  
Query Match 1.8%; Score 8; DB 22; Length 31;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
437 DDDDDDDG 444  
16 DDDDDDDG 23  
RESULT 55  
ABB34590  
ID ABB34590 standard; Peptide; 31 AA.  
XX  
AC ABB34590;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #2096 encoded by human foetal liver single exon probe.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00669.  
XX

PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483447/52.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human fetal liver -  
XX  
XX Claim 27; SEQ ID NO 27225; 639pp + sequence listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 31 AA;  
XX  
Query Match 1.8%; Score 8; DB 22; Length 31;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
437 DDDDDDDG 444  
16 DDDDDDDG 23  
RESULT 56  
ABB19996  
ID ABB19996 standard; Protein; 31 AA.  
XX  
AC ABB19996;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Protein #1995 encoded by probe for measuring heart cell gene expression.  
XX  
KW Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200157274-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00666.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX

```
XX WPI; 2001-488899/53.
DR
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 15; SEQ ID NO 21766; 530bp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 31 AA;

Query Match
Best Local Similarity 1.8%; Score 8; DB 22; Length 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444
DB 16 DDDDDDDG 23

RESULT 57
AAM55379
ID AAM55379 standard; Protein; 31 AA.
XX
AC AAM55379;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27484.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PS 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PS WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 27484; 650bp + Sequence listing; English.
XX
```

```
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 31 AA;

Query Match
Best Local Similarity 1.8%; Score 8; DB 22; Length 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444
DB 16 DDDDDDDG 23

RESULT 58
AAM67774
ID AAM67774 standard; Protein; 31 AA.
XX
AC AAM67774;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 28080.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PI 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PS WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 28080; 658bp + Sequence listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 31 AA;

Query Match
Best Local Similarity 1.8%; Score 8; DB 22; Length 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444
```



Db 16 DDDDDDDG 23

RESULT 59  
AAM15583  
ID AAM15583 standard; Protein; 31 AA.

XX AC AAM15583;

XX DT 12-OCT-2001 (first entry)

XX DE Peptide #2017 encoded by probe for measuring cervical gene expression.

XX KW Probe; human; microarray; gene expression; cervical epithelial cell;

XX KM cervical cancer.

XX OS Homo sapiens.

XX PN WO200157278-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00670.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PS analyzing gene expression in human cervical epithelial cells -

XX PS Claim 27; SEQ ID NO 20409; 487bp; English.

XX CC The present invention relates to human single exon nucleic acid probes

XX CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded

XX CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs

XX CC can be used to produce a single exon microarray, which can be used for

XX CC measuring human gene expression in a sample derived from human cervical

XX CC epithelial cells. By measuring gene expression, the probes are therefore

XX CC useful in grading and/or staging of diseases of the cervix, notably

XX CC cervical cancer.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 31 AA;

Query Match 1.8%; Score 8; DB 22; Length 31;

Best Local Similarity 100.0%; Pred. No. 6.7; 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444

Db 16 DDDDDDDG 23

RESULT 60

AAM28075  
ID AAM28075 standard; Protein; 31 AA.

XX AC AAM28075;

DT 17-OCT-2001 (first entry)

XX DE Peptide #2112 encoded by probe for measuring placental gene expression.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;

XX KM genetic disorder.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00663.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-48897/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PS analyzing gene expression in human placenta -

XX PS Claim 27; SEQ ID NO 28344; 654bp; English.

XX CC The present invention relates to single exon nucleic acid probes (SENP:

XX CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one

XX CC such probe. The probes are useful for producing a microarray for

XX CC predicting, measuring and displaying gene expression in samples derived

XX CC from human placenta. The probes are useful for antenatal diagnosis of

XX CC human genetic disorders.

XX SQ Sequence 31 AA;

Query Match 1.8%; Score 8; DB 22; Length 31;

Best Local Similarity 100.0%; Pred. No. 6.7; 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444

Db 16 DDDDDDDG 23

RESULT 61

AAM03327  
ID AAM03327 standard; Protein; 31 AA.

XX AC AAM03327;

XX DT 09-OCT-2001 (first entry)

XX DE Peptide #2009 encoded by probe for measuring breast gene expression.

XX KW Probe; human; breast disease; breast cancer; development disorder;

XX KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX OS Homo sapiens.

XX PN WO200157270-A2.

XX PD 09-AUG-2001.

XX PF 29-JAN-2001; 2001WO-US00661.

```
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-234687P.
PR 04-OCT-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
XX
PT Novel single exon nucleic acid probe used to measuring gene expression
PS in a human breast -
XX
XX
PS Claim 27; SEQ ID No 12067; 322pp; English.
XX
XX
CC The present invention relates to novel single exon nucleic acid probes
CC (See A100010-A110067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 31 AA;
XX
XX
Query Match
Best Local Similarity 1.8%; Score 8; DB 22; Length 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDDG 444
DB 16 DDDDDDDG 23
XX
XX
RESULT 62
ABG37322
ID ABG37322 standard; peptide; 31 AA.
XX
XX
ABG37322;
XX
XX
DT 19-AUG-2002 (first entry)
XX
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 26987.
XX
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200186003-A2.
XX
XX
PD 15-NOV-2001.
XX
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
XX
PR 04-FEB-2000; 2000US-180312P.
```

```
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PS measure gene expression in human lung samples -
XX
XX
PS Claim 27; SEQ ID No 26987; 634pp; English.
XX
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12614 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 31 AA;
XX
XX
Query Match
Best Local Similarity 1.8%; Score 8; DB 23; Length 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDDG 444
DB 16 DDDDDDDG 23
XX
XX
RESULT 63
AAB83381
ID AAB83381 standard; peptide; 41 AA.
```

XX AAB83381;  
AC  
XX 26-MAR-2002 (first entry)  
DT  
XX Lunasin fragment #2.  
DE  
XX Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
KM prostate; liver; kidney; therapy; soybean.  
XX  
OS Glycine max.  
XX  
PN WO200172784-A2.  
XX  
PD 04-OCT-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09453.  
XX  
PR 24-MAR-2000; 2000US-0534705.  
XX  
XX (FILG-) FILGEN BIOSCIENCES INC.  
PI Galvez AF;  
XX  
DR WPI; 2001-648438/74.  
XX  
XX Treatment or prevention of cancer by administering a lunasin peptide  
PT which binds to non-acetylated histones H3 and H4 and prevents their  
PT acetylation in hypoacetylated repressed chromatin  
XX  
PS Claim 7; Page 41; 49pp; English.  
XX  
XX This sequence represents a lunasin peptide used in the method of the  
CC invention. The method is for the treatment or prevention of cancer,  
CC comprising administration of a lunasin peptide or its active fragment or  
CC analog. The invention is used to treat or prevent cancer, particularly  
CC in a human. The cancers that can be treated and prevented include  
CC those of the colon, upper gastrointestinal tract, breast, prostate,  
CC liver, kidney or any other internal organs or tissues.  
XX  
SQ Sequence 41 AA;  
  
Query Match 1.8%; Score 8; DB 22; Length 41;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 436 GDDDDDDDD 443  
|||  
34 GDDDDDDDD 41  
  
RESULT 64  
AAB62615  
ID AAB62615 standard; peptide; 41 AA.  
XX  
AC AAB62615;  
XX  
DT 23-JUL-2001 (first entry)  
XX  
DE Soybean lunasin peptide variant (residues 1-41).  
XX  
KM Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.  
XX  
OS Glycine max.  
XX  
PN WO200134808-A2.  
XX  
PD 17-MAY-2001.  
XX  
PF 12-NOV-2000; 2000WO-US31211.  
XX  
PR 12-NOV-1999; 99US-0165334.  
XX

PA (FILG-) FILGEN BIOSCIENCES INC.  
XX  
PI Galvez AF;  
XX  
DR WPI; 2001-343605/36.  
XX  
XX Large scale production of lunasin, a cancer preventive peptide from  
PT soybean, by preparing lunasin gene constructs using Pichia expression  
PT vectors, optimizing lunasin expression parameters and purifying lunasin  
PT  
XX  
PS Claim 7; Page 60; 67pp; English.  
XX  
XX The invention provides a method of producing lunasin, a cancer preventive  
CC peptide from soybean, or its active variant by recombinant DNA technology  
CC in large quantities. The method comprises preparing lunasin gene  
CC constructs or its variants using protein expression vectors, optimizing  
CC lunasin expression parameters for large-scale production and performing  
CC a series of isolation and purification to obtain large quantities of  
CC biologically active recombinant lunasin peptide. The lunasin peptide  
CC exhibits inhibitory effects against malignant transformation of cells  
CC induced by chemical carcinogens and viral oncogenes. Sequences  
CC AAB62614-24 represent soybean lunasin variants.  
XX  
SQ Sequence 41 AA;  
  
Query Match 1.8%; Score 8; DB 22; Length 41;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 436 GDDDDDDDD 443  
|||  
Db 34 GDDDDDDDD 41  
  
RESULT 65  
AAB83380  
ID AAB83380 standard; peptide; 42 AA.  
XX  
AC AAB83380;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
XX Lunasin fragment #1.  
DE  
XX Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
KM prostate; liver; kidney; therapy; soybean.  
XX  
OS Glycine max.  
XX  
PN WO200172784-A2.  
XX  
PD 04-OCT-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09453.  
XX  
PR 24-MAR-2000; 2000US-0534705.  
XX  
XX (FILG-) FILGEN BIOSCIENCES INC.  
PI Galvez AF;  
XX  
DR WPI; 2001-648438/74.  
XX  
XX Treatment or prevention of cancer by administering a lunasin peptide  
PT which binds to non-acetylated histones H3 and H4 and prevents their  
PT acetylation in hypoacetylated repressed chromatin  
XX  
PS Claim 7; Page 40; 49pp; English.  
XX  
XX This sequence represents a lunasin peptide used in the method of the  
CC invention. The method is for the treatment or prevention of cancer,  
CC comprising administration of a lunasin peptide or its active fragment or

CC analog. The invention is used to treat or prevent cancer, particularly  
CC in a human. The cancers that can be treated and prevented include  
CC those of the colon, upper gastrointestinal tract, breast, prostate,  
CC liver, kidney or any other internal organs or tissues.  
XX  
SQ Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDDDD 443  
Db 34 GDDDDDDDD 41

RESULT 66  
ABB31264  
ABB31264 standard; Peptide; 42 AA.

ABB31264;

01-FEB-2002 (first entry)

Peptide #3915 encoded by breast cell single exon nucleic acid probe.

Human; microarray; single exon probe; gene expression; breast;

disease; cancer.

Homo sapiens.

WO200157271-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00662.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-496933/54.

New spatially-addressable set of single exon nucleic acid probes,  
useful for measuring gene expression in sample derived from human  
breast, comprises number of single exon nucleic acid probes

Claim 27; SEQ ID NO 14232; 327bp + sequence listing; English.

The invention relates to a spatially-addressable set of single exon  
nucleic acid probes for measuring gene expression in a sample derived  
from human breast and BT 474 cells. The method involves contacting  
the probes with a collection of detectably labelled nucleic acids  
derived from mRNA of human breast, and then measuring the label  
bound to each probe of the microarray. The probes are useful for  
verifying the expression of regions of genomic DNA predicted to  
encode proteins. They are useful for gene discovery, and for  
determining predisposition and/or prognosing breast disease. Gene  
expression analysis is useful for assessing the toxicity of chemical  
agents on cells. The microarray of this invention presents a far greater  
diversity of probes for measuring gene expression, with far less bias  
than expressed sequence tag microarrays. The method is suitable for  
rapid production of functional information from genomic sequence. The  
present sequence is a peptide encoded by a single exon nucleic acid  
probe of the invention.

CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDDDG 444  
Db 14 DDDDDDDDG 21

RESULT 67  
ABB36472  
ABB36472 standard; Peptide; 42 AA.

ABB36472;

04-FEB-2002 (first entry)

Peptide #3978 encoded by human foetal liver single exon probe.

Human; foetal liver; gene expression; single exon nucleic acid probe.

Homo sapiens.

WO200157277-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00669.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483447/52.

Human genome-derived single exon nucleic acid probes useful for  
analyzing gene expression in human foetal liver

Claim 27; SEQ ID NO 29107; 639bp + sequence listing; English.

The invention relates to a single exon nucleic acid probe for  
measuring human gene expression in a sample derived from human foetal  
liver. The single exon nucleic acid probes may be used for predicting,  
measuring and displaying gene expression in samples derived from human  
foetal liver. The present sequence is a peptide encoded by a single exon  
nucleic acid probe of the invention.

Note: The sequence data for this patent did not form part of the  
printed specification, but was obtained in electronic format directly  
from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDDDG 444  
Db 14 DDDDDDDDG 21



RESULT 68  
ABB21815  
ID ABB21815 standard; Protein; 42 AA.  
XX  
AC ABB21815;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Protein #3814 encoded by probe for measuring heart cell gene expression.  
XX  
KW Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200157274-A2.  
XX  
PR 09-AUG-2001.  
XX  
PR 30-JAN-2001; 2001WO-US00666.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488899/53.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts -  
XX  
PS Claim 15; SEQ ID NO 23585; 530pp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 42 AA;  
XX  
Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 437 DDDDDDDG 444  
|||  
Db 14 DDDDDDDG 21

DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29342.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00667.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483446/52.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -  
XX  
PS Example 4; SEQ ID NO: 29342; 650pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.  
XX  
SQ Sequence 42 AA;  
XX  
Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 437 DDDDDDDG 444  
|||  
Db 14 DDDDDDDG 21

RESULT 70  
AAM69636  
ID AAM69636 standard; Protein; 42 AA.  
XX  
AC AAM69636;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 29942.  
XX  
KW Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukaemia; lymphoma; myeloma.  
XX  
OS Homo sapiens.  
XX  
PN WO200157276-A2.  
XX  
PD 09-AUG-2001.  
XX

PF 30-JAN-2001; 2001WO-US00668.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488900/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human bone marrow -  
XX  
XX  
Example 4; SEQ ID NO: 29942; 658bp + Sequence Listing; English.  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention.  
XX  
SQ Sequence 42 AA;  
XX  
Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 437 DDDDDDDG 444  
DB 14 DDDDDDDG 21  
XX  
RESULT 71  
AAM17452  
ID AAM17452 standard; Protein; 42 AA.  
XX  
AC AAM17452;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
XX Peptide #3886 encoded by probe for measuring cervical gene expression.  
KM Probe; human; microarray; gene expression; cervical epithelial cell;  
KM cervical cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157278-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00670.  
XX  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488901/53.

XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
XX  
PS Claim 27; SEQ ID No 22278; 487bp; English.  
XX  
XX The present invention relates to human single exon nucleic acid probes  
CC (SENPs: see AAI10068-AA128459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 42 AA;  
XX  
Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 437 DDDDDDDG 444  
DB 14 DDDDDDDG 21  
XX  
RESULT 72  
AAM29972  
ID AAM29972 standard; Protein; 42 AA.  
XX  
AC AAM29972;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
XX Peptide #4009 encoded by probe for measuring placental gene expression.  
DE Probe; microarray; human; placenta; antenatal diagnosis;  
KM genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00663.  
XX  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488897/53.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
PS Claim 27; SEQ ID No 30241; 654bp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes (SENPs:  
CC see AAI131315-AA157546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived

CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.

XX Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
|||  
Db 14 DDDDDDDG 21

RESULT 73

AAM05129  
ID AAM05129 standard; Protein; 42 AA.

XX AAM05129;

XX 09-OCT-2001 (first entry)

XX Peptide #3811 encoded by probe for measuring breast gene expression.

XX Probe; human; breast disease; breast cancer; development disorder;  
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.

OS Homo sapiens.

PN WO200157270-A2.

PD 09-AUG-2001.

PF 29-JAN-2001; 2001WO-US00661.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression  
XX in a human breast -

PS Claim 27; SEQ ID No 13869; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes  
XX (see AII00010-AII10067). The present sequence is a peptide encoded by one  
XX such probe. The probes are useful for measuring human gene expression in  
XX a human breast sample, where the probe hybridises at high stringency to a  
XX nucleic acid expressed in the human breast. The probes are useful for  
XX predicting, diagnosing, grading, staging, monitoring and prognosing  
XX diseases of the human breast, particularly those diseases with polygenic  
XX aetiology. The diseases include: breast cancer, disorders of development,  
XX inflammatory diseases of the breast, fibrocystic changes, proliferative  
XX breast disease and non-carcinoma tumours.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.

SO Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
|||  
Db 14 DDDDDDDG 21

RESULT 74

AAB62614  
ID AAB62614 standard; peptide; 42 AA.

XX AAB62614;

XX 23-JUL-2001 (first entry)

XX Soybean lunasin peptide variant (residues 1-42).

XX Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.

OS Glycine max.

PN WO200134808-A2.

PD 17-MAY-2001.

PF 12-NOV-2000; 2000WO-US31211.

PR 12-NOV-1999; 99US-0165334.

PA (FILG-) FILGEN BIOSCIENCES INC.

PI Galvez AF;

DR WPI; 2001-343605/36.

XX Large scale production of lunasin, a cancer preventive peptide from  
XX soybean, by preparing lunasin gene constructs using Pichia expression  
XX vectors, optimizing lunasin expression parameters and purifying lunasin

PS Claim 7; Page 60; 67pp; English.

XX The invention provides a method of producing lunasin, a cancer preventive  
XX peptide from soybean, or its active variant by recombinant DNA technology  
XX in large quantities. The method comprises preparing lunasin gene  
XX constructs or its variants using protein expression vectors, optimizing  
XX lunasin expression parameters for large-scale production and performing  
XX a series of isolation and purification to obtain large quantities of  
XX biologically active recombinant lunasin peptide. The lunasin peptide  
XX exhibits inhibitory effects against malignant transformation of cells  
XX induced by chemical carcinogens and viral oncogenes. Sequences  
XX AAB62614-24 represent soybean lunasin variants.

SO Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDDG 443  
|||  
Db 34 GDDDDDDG 41

RESULT 75

ABG39258  
ID ABG39258 standard; Peptide; 42 AA.

XX ABG39258;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 28923.

Human; single exon probe; asthma; lung cancer; COPD; ILD;  
chronic obstructive pulmonary disease; interstitial lung disease;  
familial idiopathic pulmonary fibrosis; neurofibromatosis;  
tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;  
pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
primary ciliary dyskinesia; pulmonary hypertension;  
hyaline membrane disease.  
Homo sapiens.  
WO200186003-A2.  
15-NOV-2001.  
30-JAN-2001, 2001WO-US00665.  
04-FEB-2000, 2000US-180312P.  
26-MAY-2000, 2000US-207456P.  
30-JUN-2000, 2000US-0608408.  
03-AUG-2000, 2000US-0632366.  
21-SEP-2000, 2000US-234687P.  
27-SEP-2000, 2000US-236359P.  
04-OCT-2000, 2000GB-0024263.  
(MOLE-) MOLECULAR DYNAMICS INC.  
Penn SG, Hanzel DK, Chen W, Rank DR;  
WPI, 2002-114183/15.  
Spatially-addressable set of single exon nucleic acid probes, used to  
measure gene expression in human lung samples -  
Claim 27; SEQ ID No 28923; 634p; English.  
The invention relates to a spatially-addressable set of single exon  
nucleic acid probes for measuring gene expression in a sample derived  
from human lung comprising single exon nucleic acid probes having one of  
12614 nucleic acid sequences mentioned in the specification, or their  
complements or the 12387 open reading frames derived from the 12614  
probes. Also included are a microarray comprising the novel set of  
probes; the novel set of probes which hybridise at high stringency to a  
nucleic acid expressed in the human lung; measuring gene expression in a  
sample derived from human lung, comprising (a) contacting the array with  
a collection of detectably labeled nucleic acids derived from human lung  
mRNA, and (b) measuring the label detectably bound to each probe of  
the array; identifying exons in a eukaryotic genome, comprising  
(a) algorithmically predicting at least one exon from genomic sequences  
of the eukaryote; and (b) detecting specific hybridisation of detectably  
labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
having a fragment identical to the predicted exon, the probe is included  
in the above mentioned microarray; assigning exons to a single gene,  
comprising (a) identifying exons from genomic sequence by the method  
above and (b) measuring the expression of each of the exons in several  
tissues and/or cell types using hybridisation to a single exon  
microarray having a probe with the exon, where a common pattern of  
expression of the exons in the tissues and/or cell types indicates that  
the exons should be assigned to a single gene; a peptide comprising one  
of 12011 sequences, mentioned in the specification, or encoded by the  
probes/open reading frames (ORF). The probes are used for gene  
expression analysis, and for identifying exons in a gene, particularly  
using human lung derived mRNA and for the study of lung diseases  
such as asthma, lung cancer, chronic obstructive pulmonary disease  
(COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic  
pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
and hyaline membrane disease. The present sequence is a peptide/protein  
encoded by a single exon probe of the invention.

CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 42 AA;

Query Match 1.8%; Score 8; DB 23; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
|||  
Db 14 DDDDDDDG 21

Search completed: May 1, 2003, 20:53:02  
Job time : 49 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:55:38 ; Search time 48 Seconds  
(without alignments)  
817.926 Million cell updates/sec

Title: US-09-757-982-5  
Perfect score: 455  
Sequence: 1 MSSIGASFVQIKFDLQFFE.....GDDDDDDGEEEDNDMDNSE 455

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 328255 seqs, 86286685 residues  
Word size : 0

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| 1          | 455   | 100.0       | 455    | 10 | US-09-757-982-5     |
| 2          | 80    | 17.6        | 92     | 9  | US-09-764-868-799   |
| 3          | 9     | 2.0         | 502    | 10 | US-09-903-068-12    |
| 4          | 8     | 1.8         | 21     | 9  | US-10-252-256-3     |
| 5          | 8     | 1.8         | 22     | 10 | US-09-864-761-36493 |
| 6          | 8     | 1.8         | 25     | 9  | US-10-252-256-5     |
| 7          | 8     | 1.8         | 31     | 10 | US-09-864-761-35294 |
| 8          | 8     | 1.8         | 42     | 10 | US-09-864-761-37113 |
| 9          | 8     | 1.8         | 43     | 9  | US-10-252-256-1     |
| 10         | 8     | 1.8         | 49     | 9  | US-09-836-392-34    |
| 11         | 8     | 1.8         | 236    | 10 | US-09-809-545A-6    |
| 12         | 8     | 1.8         | 286    | 10 | US-09-801-368-188   |
| 13         | 8     | 1.8         | 328    | 10 | US-09-862-027-18    |
| 14         | 8     | 1.8         | 416    | 9  | US-09-291-417-5     |
| 15         | 8     | 1.8         | 416    | 10 | US-09-862-027-14    |
| 16         | 8     | 1.8         | 487    | 10 | US-09-810-808-8     |
| 17         | 8     | 1.8         | 579    | 9  | US-10-158-895-4     |
| 18         | 8     | 1.8         | 590    | 9  | US-10-158-895-15    |
| 19         | 8     | 1.8         | 847    | 9  | US-10-143-133-2     |

|    |   |     |     |    |                     |                    |
|----|---|-----|-----|----|---------------------|--------------------|
| 20 | 7 | 1.5 | 8   | 9  | US-09-882-291-22    | Sequence 22, Appl  |
| 21 | 7 | 1.5 | 9   | 9  | US-09-882-291-29    | Sequence 29, Appl  |
| 22 | 7 | 1.5 | 12  | 9  | US-09-876-904A-173  | Sequence 173, App  |
| 23 | 7 | 1.5 | 14  | 9  | US-09-911-569-80    | Sequence 80, Appl  |
| 24 | 7 | 1.5 | 20  | 10 | US-09-864-761-38908 | Sequence 38908, A  |
| 25 | 7 | 1.5 | 23  | 10 | US-09-864-761-36028 | Sequence 36028, A  |
| 26 | 7 | 1.5 | 24  | 10 | US-09-864-761-34620 | Sequence 34620, A  |
| 27 | 7 | 1.5 | 24  | 10 | US-09-864-761-36142 | Sequence 36142, A  |
| 28 | 7 | 1.5 | 26  | 10 | US-09-864-761-35447 | Sequence 35447, A  |
| 29 | 7 | 1.5 | 27  | 10 | US-09-864-761-39392 | Sequence 39392, A  |
| 30 | 7 | 1.5 | 30  | 9  | US-09-911-569-87    | Sequence 87, Appl  |
| 31 | 7 | 1.5 | 31  | 10 | US-09-864-761-33555 | Sequence 33555, A  |
| 32 | 7 | 1.5 | 34  | 10 | US-09-864-761-44046 | Sequence 44046, A  |
| 33 | 7 | 1.5 | 36  | 10 | US-09-864-761-35933 | Sequence 35933, A  |
| 34 | 7 | 1.5 | 37  | 10 | US-09-864-761-43116 | Sequence 43116, A  |
| 35 | 7 | 1.5 | 40  | 9  | US-10-252-256-2     | Sequence 2, Appl   |
| 36 | 7 | 1.5 | 40  | 10 | US-09-864-761-39102 | Sequence 39102, A  |
| 37 | 7 | 1.5 | 42  | 10 | US-09-864-761-35018 | Sequence 35018, A  |
| 38 | 7 | 1.5 | 42  | 10 | US-09-864-761-35629 | Sequence 35629, A  |
| 39 | 7 | 1.5 | 51  | 10 | US-09-864-761-39869 | Sequence 39869, A  |
| 40 | 7 | 1.5 | 61  | 10 | US-09-864-761-39079 | Sequence 39079, A  |
| 41 | 7 | 1.5 | 86  | 12 | US-10-028-247-2     | Sequence 2, Appl   |
| 42 | 7 | 1.5 | 91  | 10 | US-09-873-880-4     | Sequence 4, Appl   |
| 43 | 7 | 1.5 | 102 | 10 | US-09-864-761-44209 | Sequence 44209, A  |
| 44 | 7 | 1.5 | 110 | 10 | US-09-864-761-42215 | Sequence 42215, A  |
| 45 | 7 | 1.5 | 154 | 9  | US-10-101-464A-662  | Sequence 662, App  |
| 46 | 7 | 1.5 | 154 | 10 | US-09-864-761-33375 | Sequence 33375, A  |
| 47 | 7 | 1.5 | 161 | 10 | US-09-864-761-43542 | Sequence 43542, A  |
| 48 | 7 | 1.5 | 163 | 9  | US-10-101-464A-681  | Sequence 681, App  |
| 49 | 7 | 1.5 | 170 | 9  | US-09-764-868-816   | Sequence 816, App  |
| 50 | 7 | 1.5 | 170 | 9  | US-09-764-868-845   | Sequence 845, App  |
| 51 | 7 | 1.5 | 170 | 9  | US-09-764-891-4540  | Sequence 4540, App |
| 52 | 7 | 1.5 | 174 | 9  | US-09-764-868-633   | Sequence 633, App  |
| 53 | 7 | 1.5 | 177 | 9  | US-09-764-868-833   | Sequence 833, App  |
| 54 | 7 | 1.5 | 177 | 9  | US-09-764-891-5028  | Sequence 5028, App |
| 55 | 7 | 1.5 | 188 | 9  | US-10-101-464A-540  | Sequence 540, App  |
| 56 | 7 | 1.5 | 207 | 9  | US-09-738-626-4510  | Sequence 4510, App |
| 57 | 7 | 1.5 | 238 | 9  | US-10-101-464A-702  | Sequence 702, App  |
| 58 | 7 | 1.5 | 244 | 10 | US-09-925-301-1424  | Sequence 1424, App |
| 59 | 7 | 1.5 | 250 | 9  | US-09-764-868-1063  | Sequence 1063, App |
| 60 | 7 | 1.5 | 250 | 9  | US-09-955-999-77    | Sequence 77, Appl  |
| 61 | 7 | 1.5 | 256 | 9  | US-09-898-837A-32   | Sequence 32, Appl  |
| 62 | 7 | 1.5 | 257 | 9  | US-09-949-842-14    | Sequence 14, Appl  |
| 63 | 7 | 1.5 | 265 | 9  | US-09-764-868-821   | Sequence 821, App  |
| 64 | 7 | 1.5 | 271 | 10 | US-09-840-704-6     | Sequence 6, Appl   |
| 65 | 7 | 1.5 | 275 | 9  | US-10-274-409-5     | Sequence 5, Appl   |
| 66 | 7 | 1.5 | 277 | 9  | US-10-274-409-6     | Sequence 6, Appl   |
| 67 | 7 | 1.5 | 277 | 9  | US-10-274-409-7     | Sequence 7, Appl   |
| 68 | 7 | 1.5 | 280 | 8  | US-08-910-386A-14   | Sequence 14, Appl  |
| 69 | 7 | 1.5 | 284 | 10 | US-09-771-161A-106  | Sequence 106, App  |
| 70 | 7 | 1.5 | 295 | 9  | US-10-060-019-22    | Sequence 22, Appl  |
| 71 | 7 | 1.5 | 303 | 9  | US-09-991-211-5     | Sequence 5, Appl   |
| 72 | 7 | 1.5 | 304 | 10 | US-09-866-582-27    | Sequence 27, Appl  |
| 73 | 7 | 1.5 | 324 | 10 | US-09-746-801A-61   | Sequence 61, Appl  |
| 74 | 7 | 1.5 | 325 | 10 | US-09-815-242-5783  | Sequence 5783, App |
| 75 | 7 | 1.5 | 325 | 10 | US-09-815-242-12693 | Sequence 12693, A  |
| 76 | 7 | 1.5 | 325 | 10 | US-09-815-242-12779 | Sequence 12779, A  |
| 77 | 7 | 1.5 | 325 | 10 | US-09-815-242-13070 | Sequence 13070, A  |
| 78 | 7 | 1.5 | 329 | 10 | US-09-925-300-1268  | Sequence 1268, App |
| 79 | 7 | 1.5 | 353 | 10 | US-09-801-368-116   | Sequence 116, App  |
| 80 | 7 | 1.5 | 367 | 9  | US-10-026-021-6     | Sequence 6, Appl   |
| 81 | 7 | 1.5 | 379 | 9  | US-09-944-413-2     | Sequence 2, Appl   |
| 82 | 7 | 1.5 | 379 | 9  | US-09-944-403-2     | Sequence 2, Appl   |
| 83 | 7 | 1.5 | 379 | 9  | US-09-944-896-2     | Sequence 2, Appl   |
| 84 | 7 | 1.5 | 379 | 9  | US-09-944-944-2     | Sequence 2, Appl   |
| 85 | 7 | 1.5 | 379 | 9  | US-09-944-907-2     | Sequence 2, Appl   |
| 86 | 7 | 1.5 | 379 | 9  | US-09-944-929-2     | Sequence 2, Appl   |
| 87 | 7 | 1.5 | 379 | 9  | US-10-028-072-328   | Sequence 328, App  |
| 88 | 7 | 1.5 | 379 | 9  | US-10-121-049-328   | Sequence 328, App  |
| 89 | 7 | 1.5 | 379 | 9  | US-10-123-904-328   | Sequence 328, App  |
| 90 | 7 | 1.5 | 379 | 9  | US-10-140-470-328   | Sequence 328, App  |
| 91 | 7 | 1.5 | 379 | 9  | US-10-175-746-328   | Sequence 328, App  |
| 92 | 7 | 1.5 | 379 | 9  | US-10-176-918-328   | Sequence 328, App  |

|     |   |     |     |   |                   |                   |
|-----|---|-----|-----|---|-------------------|-------------------|
| 93  | 7 | 1.5 | 379 | 9 | US-10-176-921-328 | Sequence 328, App |
| 94  | 7 | 1.5 | 379 | 9 | US-10-137-865-328 | Sequence 328, App |
| 95  | 7 | 1.5 | 379 | 9 | US-10-140-474-328 | Sequence 328, App |
| 96  | 7 | 1.5 | 379 | 9 | US-10-142-431-328 | Sequence 328, App |
| 97  | 7 | 1.5 | 379 | 9 | US-10-143-114-328 | Sequence 328, App |
| 98  | 7 | 1.5 | 379 | 9 | US-10-140-002-328 | Sequence 328, App |
| 99  | 7 | 1.5 | 379 | 9 | US-10-142-419-328 | Sequence 328, App |
| 100 | 7 | 1.5 | 379 | 9 | US-10-123-262-328 | Sequence 328, App |

ALIGNMENTS

RESULT 1  
US-09-757-982-5  
; Sequence 5, Application US/09757982  
; Patent No. US20020094559A1  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/757,982  
; PRIOR FILING DATE: 2001-01-10  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-757-982-5

|  |     |                         |                       |                         |                    |     |
|--|-----|-------------------------|-----------------------|-------------------------|--------------------|-----|
| Query Match  |     |                         |                       |                         |                    |     |
| Best Local Similarity 100.0%; Score 455; DB 10; Length 455;  |     |                         |                       |                         |                    |     |
| Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |     |                         |                       |                         |                    |     |
| QY   | 1   | MSSLSGASFVOIKPDDLQFFENC | GGSGSVYRAKWSQDKEVA    | VKKLKEAEILSVL           | 60                 |     |
| DB   | 1   | MSSLSGASFVOIKPDDLQFFENC | GGSGSVYRAKWSQDKEVA    | VKKLKEAEILSVL           | 60                 |     |
| QY   | 61  | SHRNIIQFYGVILEPPNYGIVTE | YASLSGLYDINSNRSEEMDM  | HIMTWATD                | 120                |     |
| DB   | 61  | SHRNIIQFYGVILEPPNYGIVTE | YASLSGLYDINSNRSEEMDM  | HIMTWATD                | 120                |     |
| QY   | 121 | LHMEAPVKVIHRDLKSRNVVIA  | ADGVLCIDFGASRFHNHTHMS | LVGTFPWMAPEVIQS         | 180                |     |
| DB   | 121 | LHMEAPVKVIHRDLKSRNVVIA  | ADGVLCIDFGASRFHNHTHMS | LVGTFPWMAPEVIQS         | 180                |     |
| QY   | 181 | LPVSETCDITYSYGVVLWEMLT  | REVPFKGLEGLQVAVLVEKNE | RLTIPSSCPRSFAELH        | 240                |     |
| DB   | 181 | LPVSETCDITYSYGVVLWEMLT  | REVPFKGLEGLQVAVLVEKNE | RLTIPSSCPRSFAELH        | 240                |     |
| QY   | 241 | QCWEADAKKRPSPFKQIISILE  | SMNSDTSLPDKCN         | SFLHNKAEMWRCIEATLERL    | 300                |     |
| DB   | 241 | QCWEADAKKRPSPFKQIISILE  | SMNSDTSLPDKCN         | SFLHNKAEMWRCIEATLERL    | 300                |     |
| QY   | 301 | LSFKEQELKERERRLLKMEQKL  | TEQSNTPLLPLAR         | MSEESYFESKTEESNSAEMSCOI | 360                |     |
| DB   | 301 | LSFKEQELKERERRLLKMEQKL  | TEQSNTPLLPLAR         | MSEESYFESKTEESNSAEMSCOI | 360                |     |
| QY   | 361 | TATSNGEHGHNPSLQAMLMGFG  | DI                    | FSMNKAGAVMHS            | GMQINMOAKONSSKTSKR | 420 |
| DB   | 361 | TATSNGEHGHNPSLQAMLMGFG  | DI                    | FSMNKAGAVMHS            | GMQINMOAKONSSKTSKR | 420 |
| QY   | 421 | KVNMAFGSDPDLSEGD        | DDDDDDGEEEDNDMDNSE    | 455                     |                    |     |
| DB   | 421 | KVNMAFGSDPDLSEGD        | DDDDDDGEEEDNDMDNSE    | 455                     |                    |     |

RESULT 2  
US-09-764-868-799  
; Sequence 799, Application US/09764868

Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; PRIOR FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 799  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (4)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-868-799

|   |     |                         |                       |                |     |  |
|---|-----|-------------------------|-----------------------|----------------|-----|--|
| Query Match   |     |                         |                       |                |     |  |
| Best Local Similarity 17.6%; Score 80; DB 9; Length 92;     |     |                         |                       |                |     |  |
| Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |     |                         |                       |                |     |  |
| QY  | 54  | AEILSVLSHRNIIQFYGVILEPP | NYGIVTEYASLSGLYDINSNR | SEEMDMHIMTWATD | 113 |  |
| DB  | 13  | AEILSVLSHRNIIQFYGVILEPP | NYGIVTEYASLSGLYDINSNR | SEEMDMHIMTWATD | 72  |  |
| QY  | 114 | VAKGMHYLHMEAPVKVIHRD    | 133                   |                |     |  |
| DB  | 73  | VAKGMHYLHMEAPVKVIHRD    | 92                    |                |     |  |

RESULT 3  
US-09-903-068-12  
; Sequence 12, Application US/09903068  
; Patent No. US20020123139A1  
; GENERAL INFORMATION:  
; APPLICANT: Miyazono, Kohel; DiJke, Peter Ten;  
; Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik  
; TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins  
; Having Serine Threonine Kinase Domains And Their Use  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/903,068  
; FILING DATE: 11-Jul-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/679,187  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: PCT/GB93/02367  
; FILING DATE: 17-NO. US20020123139A1ember-1993  
; APPLICATION NUMBER: 9224057.1  
; FILING DATE: 17-NO. US20020123139A1ember-1992  
; APPLICATION NUMBER: 9304677.9  
; FILING DATE: 8-March-1993  
; APPLICATION NUMBER: 9304680.3  
; FILING DATE: 8-March-1993  
; APPLICATION NUMBER: 9311047.6  
; FILING DATE: 28-May-1993  
; APPLICATION NUMBER: 9313763.6  
; FILING DATE: 2-July-1993  
; APPLICATION NUMBER: 9136099.2

FILING DATE: 3-August-1993  
APPLICATION NUMBER: 9321344.5  
FILING DATE: 15-October-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohlei, Vineet  
REGISTRATION NUMBER: 37,003  
REFERENCE/DOCKET NUMBER: LUD 5298  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-903-068-12

Query Match 2.0%; Score 9; DB 10; Length 502;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
|||  
Db 327 HRDLKSRNV 335

## RESULT 4

US-10-252-256-3  
Sequence 3, Application US/10252256  
Publication No. US2003002765A1  
GENERAL INFORMATION:  
APPLICANT: GALVEZ, ALFREDO F.  
TITLE OF INVENTION: THERAPEUTIC PEPTIDES HAVING A MOTIF THAT BINDS  
SPECIFICALLY TO NON-ACETYLATED H3 AND H4 HISTONES FOR  
CANCER THERAPY  
FILE REFERENCE: 3729.02  
CURRENT APPLICATION NUMBER: US/10/252,256  
CURRENT FILING DATE: 2002-09-23  
PRIOR APPLICATION NUMBER: 60/165,334  
PRIOR FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 21  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: trlunasin  
US-10-252-256-3

Query Match 1.8%; Score 8; DB 9; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
|||  
Db 12 GDDDDDD 19

## RESULT 5

US-09-864-761-36493  
Sequence 36493, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aecm1ca-X-1

CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 36493  
LENGTH: 22  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC011235.2  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.97  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1  
US-09-864-761-36493

Query Match 1.8%; Score 8; DB 10; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
|||  
Db 7 DDDDDDDG 14

## RESULT 6

US-10-252-256-5  
Sequence 5, Application US/10252256  
Publication No. US2003002765A1  
GENERAL INFORMATION:  
APPLICANT: GALVEZ, ALFREDO F.



;; TITLE OF INVENTION: THERAPEUTIC PEPTIDES HAVING A MOTIF THAT BINDS  
;; TITLE OF INVENTION: SPECIFICALLY TO NON-ACETYLATED H3 AND H4 HISTONES FOR  
;; TITLE OF INVENTION: CANCER THERAPY  
;; FILE REFERENCE: 3729.02  
;; CURRENT APPLICATION NUMBER: US/10/252,256  
;; CURRENT FILING DATE: 2002-09-23  
;; PRIOR APPLICATION NUMBER: 60/165,334  
;; PRIOR FILING DATE: 1999-11-12  
;; NUMBER OF SEQ ID NOS: 15  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 5  
;; LENGTH: 25  
;; TYPE: PRT  
;; ORGANISM: Glycine max  
;; FEATURE:  
;; OTHER INFORMATION: NLS-ctrlunasin  
US-10-252-256-5

Query Match 1.8%; Score 8; DB 9; Length 25;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
Db 16 GDDDDDD 23

RESULT 7  
US-09-864-761-35294  
; Sequence 35294, Application US/09864761  
; Patent No. US20020048763A1

;; GENERAL INFORMATION:

;; APPLICANT: Penn, Sharon G.

;; APPLICANT: Rank, David R.

;; APPLICANT: Hanzel, David K.

;; APPLICANT: Chen, Wensheng

;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

;; FILE REFERENCE: Aeomica-X-1

;; CURRENT APPLICATION NUMBER: US/09/864,761

;; CURRENT FILING DATE: 2001-05-23

;; PRIOR APPLICATION NUMBER: US 60/180,312

;; PRIOR FILING DATE: 2000-02-04

;; PRIOR APPLICATION NUMBER: US 60/207,456

;; PRIOR FILING DATE: 2000-05-26

;; PRIOR APPLICATION NUMBER: US 09/632,366

;; PRIOR FILING DATE: 2000-08-03

;; PRIOR APPLICATION NUMBER: GB 24263.6

;; PRIOR FILING DATE: 2000-10-04

;; PRIOR APPLICATION NUMBER: US 60/236,359

;; PRIOR FILING DATE: 2000-09-27

;; PRIOR APPLICATION NUMBER: PCT/US01/00666

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00667

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00664

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00669

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00665

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00668

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00663

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00662

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00661

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00670

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: US 60/234,687

;; PRIOR FILING DATE: 2000-09-21

;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 35294  
;; LENGTH: 31  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens

;; FEATURE:

;; OTHER INFORMATION: MAP TO AC011235.1

;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3

;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.6

;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6

;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.93

;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7

;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2

;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1

;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6

;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4

;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6

US-09-864-761-35294

Query Match 1.8%; Score 8; DB 10; Length 31;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
Db 16 DDDDDDDG 23

RESULT 8  
US-09-864-761-37113  
; Sequence 37113, Application US/09864761  
; Patent No. US20020048763A1

;; GENERAL INFORMATION:

;; APPLICANT: Penn, Sharon G.

;; APPLICANT: Rank, David R.

;; APPLICANT: Hanzel, David K.

;; APPLICANT: Chen, Wensheng

;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

;; FILE REFERENCE: Aeomica-X-1

;; CURRENT APPLICATION NUMBER: US/09/864,761

;; CURRENT FILING DATE: 2001-05-23

;; PRIOR APPLICATION NUMBER: US 60/180,312

;; PRIOR FILING DATE: 2000-02-04

;; PRIOR APPLICATION NUMBER: US 60/207,456

;; PRIOR FILING DATE: 2000-05-26

;; PRIOR APPLICATION NUMBER: US 09/632,366

;; PRIOR FILING DATE: 2000-08-03

;; PRIOR APPLICATION NUMBER: GB 24263.6

;; PRIOR FILING DATE: 2000-10-04

;; PRIOR APPLICATION NUMBER: US 60/236,359

;; PRIOR FILING DATE: 2000-09-27

;; PRIOR APPLICATION NUMBER: PCT/US01/00666

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00667

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00664

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00669

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00665

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00668

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00663

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00662

;; PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37113
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AJ229043.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 9.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 8.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 9.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 8.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.9
; OTHER INFORMATION: EST_HUMAN HIT: AW63068.1, EVALUE 3.00e-07
US-09-864-761-37113
```

```
Query Match
Best Local Similarity 1.8%; Score 8; DB 10; Length 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 437 DDDDDDDG 444
Db 14 DDDDDDDG 21
```

```
RESULT 9
US-10-252-256-1
```

```
; Sequence 1, Application US/10252256
; Publication No. US2003002765A1
; GENERAL INFORMATION:
; APPLICANT: GALVEZ, ALFREDO F.
; TITLE OF INVENTION: THERAPEUTIC PEPTIDES HAVING A MOTIF THAT BINDS
; TITLE OF INVENTION: SPECIFICALLY TO NON-ACETYLATED H3 AND H4 HISTONES FOR
; FILE REFERENCE: 3729.02
; CURRENT APPLICATION NUMBER: US/10/252,256
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 60/165,334
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Lunasin
US-10-252-256-1
```

```
Query Match
Best Local Similarity 1.8%; Score 8; DB 9; Length 43;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 436 GDDDDDDD 443
Db 34 GDDDDDDD 41
```

```
RESULT 10
US-09-836-392-34
; Sequence 34, Application US/09836392
; Patent No. US20020173458A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides,
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: PT020P1
; CURRENT APPLICATION NUMBER: US/09/836,392
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/165,914
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/189,027
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-392-34
```

```
Query Match
Best Local Similarity 1.8%; Score 8; DB 9; Length 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 146 VLKICDFG 153
Db 30 VLKICDFG 37
```

```
RESULT 11
```

```
US-09-809-545A-6
; Sequence 6, Application US/09809545A
; Patent No. US20020110804A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence W.
; APPLICANT: White, R. Tyler
; TITLE OF INVENTION: SECRETED FACTORS
; FILE REFERENCE: SCIOS.017A
; CURRENT APPLICATION NUMBER: US/09/809,545A
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-809-545A-6
```

```
Query Match
Best Local Similarity 1.8%; Score 8; DB 10; Length 236;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 436 GDDDDDDD 443
Db 41 GDDDDDDD 48
```

```
RESULT 12
```

```
US-09-801-368-188
; Sequence 188, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
```

APPLICANT: Madden, Kevin  
APPLICANT: Maxon, Mary  
APPLICANT: Milne, Todd  
APPLICANT: No. US20020128250A1man, Thea  
APPLICANT: Royer, John  
APPLICANT: Salama, Sofie  
APPLICANT: Sherman, Amir  
APPLICANT: Silva, Jeff  
APPLICANT: Summers, Eric  
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
FILE REFERENCE: 109272.147  
CURRENT APPLICATION NUMBER: US/09/801,368  
CURRENT FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 09/487,558  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: US 60/160,587  
PRIOR FILING DATE: 1999-10-20  
NUMBER OF SEQ ID NOS: 440  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 188  
LENGTH: 286  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-801-368-188

Query Match 1.8%; Score 8; DB 10; Length 286;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDDDD 443  
DB 108 GDDDDDDDD 115

RESULT 13  
US-09-862-027-18  
Sequence 18, Application US/09862027  
Patent No. US20020142428A1  
GENERAL INFORMATION:  
APPLICANT: Hodge, Martin R.  
TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof  
FILE REFERENCE: 35800/234862  
CURRENT APPLICATION NUMBER: US/09/862,027  
CURRENT FILING DATE: 2001-05-21  
PRIOR APPLICATION NUMBER: US 09/345,473  
PRIOR FILING DATE: 1999-06-30  
NUMBER OF SEQ ID NOS: 82  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 18  
LENGTH: 328  
TYPE: PRT  
ORGANISM: C. elegans  
US-09-862-027-18

Query Match 1.8%; Score 8; DB 10; Length 328;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 51 EKEAEILS 58  
DB 90 EKEAEILS 97

RESULT 14  
US-09-291-417-5  
Sequence 5, Application US/09291417A  
Publication No. US20030050230A1  
GENERAL INFORMATION:  
APPLICANT: PLOWMAN, GREGORY  
APPLICANT: MARTINEZ, RICARDO  
APPLICANT: WHITE, DAVID  
TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
FILE REFERENCE: 240/300

CURRENT APPLICATION NUMBER: US/09/291,417A  
CURRENT FILING DATE: 1999-04-13  
EARLIER APPLICATION NUMBER: US 60/081,784  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 147  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 5  
LENGTH: 416  
TYPE: PRT  
ORGANISM: Mammalian (Human) STLK2  
US-09-291-417-5

Query Match 1.8%; Score 8; DB 9; Length 416;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 172 WMAPEVIQ 179  
DB 185 WMAPEVIQ 192

RESULT 15  
US-09-862-027-14  
Sequence 14, Application US/09862027  
Patent No. US20020142428A1  
GENERAL INFORMATION:  
APPLICANT: Hodge, Martin R.  
TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof  
FILE REFERENCE: 35800/234862  
CURRENT APPLICATION NUMBER: US/09/862,027  
CURRENT FILING DATE: 2001-05-21  
PRIOR APPLICATION NUMBER: US 09/345,473  
PRIOR FILING DATE: 1999-06-30  
NUMBER OF SEQ ID NOS: 82  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14  
LENGTH: 416  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-862-027-14

Query Match 1.8%; Score 8; DB 10; Length 416;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 172 WMAPEVIQ 179  
DB 185 WMAPEVIQ 192

RESULT 16  
US-09-810-808-8  
Sequence 8, Application US/09810808  
Patent No. US20020042114A1  
GENERAL INFORMATION:  
APPLICANT: Au-Young, Janice  
Guegler, Karl J.  
Hawkins, Phillip R.  
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/810,808  
FILING DATE: 15-Mar-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/541,228  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0118 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1117791  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-810-808-8

Query Match 1.8%; Score 8; DB 10; Length 487;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
Db 190 WMAPEVIQ 197

RESULT 17  
US-10-158-895-4  
Sequence 4, Application US/10158895  
Patent No. US20020155624A1  
GENERAL INFORMATION:  
APPLICANT: ONO, KOICHIRO  
APPLICANT: OHTOMO, TOSHIHIKO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/10/158,895  
CURRENT FILING DATE: 2002-06-03  
PRIOR APPLICATION NUMBER: US/09/529,279  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188  
PRIOR FILING DATE: 1997-10-22  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 579  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-158-895-4

Query Match 1.8%; Score 8; DB 9; Length 579;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
Db 170 VLKICDFG 177

RESULT 18  
US-10-158-895-15  
Sequence 15, Application US/10158895  
Patent No. US20020155624A1

GENERAL INFORMATION:  
APPLICANT: ONO, KOICHIRO  
APPLICANT: OHTOMO, TOSHIHIKO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/10/158,895  
CURRENT FILING DATE: 2002-06-03  
PRIOR APPLICATION NUMBER: US/09/529,279  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188  
PRIOR FILING DATE: 1997-10-22  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 15  
LENGTH: 590  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-158-895-15

Query Match 1.8%; Score 8; DB 9; Length 590;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
Db 170 VLKICDFG 177

RESULT 19  
US-10-143-133-2  
Sequence 2, Application US/10143133  
Publication No. US20020197658A1  
GENERAL INFORMATION:  
APPLICANT: Yoganathan, Thillainathan  
APPLICANT: Delaney, Allen  
TITLE OF INVENTION: Cancer Associated Protein Kinase and Its Use  
FILE REFERENCE: KINE-023  
CURRENT APPLICATION NUMBER: US/10/143,133  
CURRENT FILING DATE: 2002-05-09  
PRIOR APPLICATION NUMBER: 60/290,555  
PRIOR FILING DATE: 2001-05-10  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 847  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-143-133-2

Query Match 1.8%; Score 8; DB 9; Length 847;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRLDKS 136  
Db 237 VIHRLDKS 244

RESULT 20  
US-09-882-291-22  
Sequence 22, Application US/09882291  
Publication No. US20030040472A1  
GENERAL INFORMATION:  
APPLICANT: Zealand Pharmaceuticals A/S  
TITLE OF INVENTION: No. US20030040472A1 Peptide Conjugates  
FILE REFERENCE: 007-2001  
CURRENT APPLICATION NUMBER: US/09/882,291  
CURRENT FILING DATE: 2001-06-15  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: PatentIn version 3.1



; SEQ ID NO 22  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide sequence  
US-09-882-291-22

Query Match 1.5%; Score 7; DB 9; Length 8;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 2 DDDDDDD 8

RESULT 21

US-09-882-291-29  
Sequence 29, Application US/09882291  
Publication No. US20030040472A1

GENERAL INFORMATION:

APPLICANT: Zealand Pharmaceuticals A/S  
TITLE OF INVENTION: No. US20030040472A1e1 Peptide Conjugates

FILE REFERENCE: 007-2001

CURRENT APPLICATION NUMBER: US/09/882,291

CURRENT FILING DATE: 2001-06-15

NUMBER OF SEQ ID NOS: 77

SOFTWARE: PatentIn version 3.1

SEQ ID NO 29

LENGTH: 9

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide sequence

US-09-882-291-29

Query Match

Best Local Similarity 1.5%; Score 7; DB 9; Length 9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443

Db 3 DDDDDDD 9

RESULT 22

US-09-876-904A-173

Sequence 173, Application US/09876904A  
Publication No. US20030072794A1

GENERAL INFORMATION:

APPLICANT: BOULIKAS, TENI

TITLE OF INVENTION: ENCAPSULATION OF PLASMID DNA (LIPOGENES TM) AND THERAPEUTIC

TITLE OF INVENTION: AGENTS WITH NUCLEAR LOCALIZATION SIGNAL/FUSOGENIC PEPTIDE

TITLE OF INVENTION: CONJUGATES INTO TARGETED LIPOSOME COMPLEXES

FILE REFERENCE: TB-2002.00

CURRENT APPLICATION NUMBER: US/09/876,904A

CURRENT FILING DATE: 2001-06-08

PRIOR APPLICATION NUMBER: US 60/210,925

PRIOR FILING DATE: 2000-06-09

NUMBER OF SEQ ID NOS: 629

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 173

LENGTH: 12

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Karyophilic peptide

US-09-876-904A-173

Query Match

Best Local Similarity 1.5%; Score 7; DB 9; Length 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRLDK 135  
Db 5 VIHRLDK 11

RESULT 23

US-09-911-569-80

Sequence 80, Application US/09911569  
Publication No. US20030069173A1

GENERAL INFORMATION:

APPLICANT: HAWLEY-NELSON, PAMELA

LAN, JIANQING

SHIH, POJEN

JESSE, JOEL A.

SCHIFFERLI, KEVIN P.

GEBEYEHU, GUILIAT

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: GREENLEE, WINNER & SULLIVAN

STREET: 5370 MANHATTAN CIRCLE, SUITE 201

CITY: BOULDER

STATE: CO

COUNTRY: US

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/911,569

FILING DATE: 23-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/039,780

FILING DATE: 16-MAR-1998

ATTORNEY/AGENT INFORMATION:

NAME: SULLIVAN, SALLY A.

REGISTRATION NUMBER: 32,064

REFERENCE/DOCKET NUMBER: 32-95D

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303)499-8080

TELEFAX: (303)499-8089

INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: Modified-site

LOCATION: 14

OTHER INFORMATION: /product= "OTHER"

/note= "G AT POSITION 14 CAN BE ABSENT"

SEQUENCE DESCRIPTION: SEQ ID NO: 80:

US-09-911-569-80

Query Match

Best Local Similarity 1.5%; Score 7; DB 9; Length 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 DDDDDDG 444

Db 8 DDDDDDG 14

RESULT 24

US-09-864-761-38908

```
Sequence 38908, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 38908
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004633.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
US-09-864-761-38908

Query Match 1.5%; Score 7; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 DDDDDDD 443
Db 13 DDDDDDD 19
```

```
RESULT 25
US-09-864-761-36028
Sequence 36028, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36028
LENGTH: 23
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC011309.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.3e+02
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
US-09-864-761-36028

Query Match 1.5%; Score 7; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 22;
```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443

Db 3 DDDDDDD 9

## RESULT 26

US-09-864-761-34620

; Sequence 34620, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 34620

; LENGTH: 24

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC009491.2; SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3

; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3

; US-09-864-761-34620

Query Match 1.5%; Score 7; DB 10; Length 24;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443

Db 5 DDDDDDD 11

## RESULT 27

US-09-864-761-36142

; Sequence 36142, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 36142

; LENGTH: 24

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC007159.4; SIGNAL = 3.3

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 13

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 17

; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 17

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2  
US-09-864-761-36142

Query Match 1.5%; Score 7; DB 10; Length 24;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDD 443  
Db 1 DDDDDD 7

## RESULT 28

Sequence 35447, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wenheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 35447  
LENGTH: 26  
TYPE: PRT

ORGANISM: Homo sapiens  
FEATURE:

OTHER INFORMATION: MAP TO AL031661.16, SIGNAL = 2.5  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7  
US-09-864-761-35447

Query Match 1.5%; Score 7; DB 10; Length 26;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443

Db 3 DDDDDD 9

## RESULT 29

Sequence 39392, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wenheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29



NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 39392  
LENGTH: 27  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC005959.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN HEILA, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN RETAL, LIVER, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7  
OTHER INFORMATION: EST\_HUMAN HIT: T95763.1, EVALUUE 2.10e+00  
US-09-864-761-39392

Query Match 1.5%; Score 7; DB 10; Length 27;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 3 DDDDDDD 9

RESULT 30  
US-09-911-569-87

Sequence 87, Application US/09911569  
Publication No. US20030069173A1

GENERAL INFORMATION:

APPLICANT: HAWLEY-NELSON, PAMELA

LAN, JIANQING

SHIH, POJEN

JESSE, JOEL A.

SCHIFFERLI, KEVIN P.

GEBEYEHU, GULILAT

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: GREENLEE, WINNER & SULLIVAN

STREET: 5370 MANHATTAN CIRCLE, SUITE 201

CITY: BOULDER

STATE: CO

COUNTRY: US

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/09/911,569

FILING DATE: 23-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/039,780

FILING DATE: 16-MAR-1998

ATTORNEY/AGENT INFORMATION:

NAME: SULLIVAN, SALLY A.

REGISTRATION NUMBER: 32,064

REFERENCE/DOCKET NUMBER: 32-95D

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303)499-8080

TELEFAX: (303)499-8089

INFORMATION FOR SEQ ID NO: 87:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 30  
OTHER INFORMATION: /product= "OTHER"  
/note= "G AT POSITION 30 CAN BE ABSENT"  
SEQUENCE DESCRIPTION: SEQ ID NO: 87:  
US-09-911-569-87

Query Match 1.5%; Score 7; DB 9; Length 30;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 DDDDDDG 444  
Db 24 DDDDDDG 30

RESULT 31

US-09-864-761-33555

Sequence 33555, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aecomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

FILE REFERENCE: Aecomica-X-1

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 33555

LENGTH: 31  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AP000032.1  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.4  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.1  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.6  
US-09-864-761-33555

Query Match 1.5%; Score 7; DB 10; Length 31;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDD 443  
DB 5 DDDDDDD 11

## RESULT 32

US-09-864-761-44046  
Sequence 44046, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 44046  
LENGTH: 34  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC002478.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.66  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.74  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.73  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.88  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.65  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.78  
US-09-864-761-44046

Query Match 1.5%; Score 7; DB 10; Length 34;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDD 443  
DB 4 DDDDDDD 10

## RESULT 33

US-09-864-761-35933  
Sequence 35933, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687

```
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 35933
;; LENGTH: 36
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC010792.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
US-09-864-761-35933

Query Match      1.5%; Score 7; DB 10; Length 36;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
Db 9 DDDDDDD 15

RESULT 34
US-09-864-761-43116
; Sequence 43116, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
```

```
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 43116
;; LENGTH: 37
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL008631.1
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.95
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
US-09-864-761-43116

Query Match      1.5%; Score 7; DB 10; Length 37;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
Db 3 DDDDDDD 9

RESULT 35
US-10-252-256-2
; Sequence 2, Application US/10252256
; Publication No. US2003002765A1
; GENERAL INFORMATION:
; APPLICANT: GALVEZ, ALFREDO F.
; TITLE OF INVENTION: THERAPEUTIC PEPTIDES HAVING A MOTIF THAT BINDS
; TITLE OF INVENTION: SPECIFICALLY TO NON-ACETYLATED H3 AND H4 HISTONES FOR
; FILE REFERENCE: 3729.02
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US/10/252,256
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 60/165,334
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Lunasin-GRG
US-10-252-256-2

Query Match      1.5%; Score 7; DB 9; Length 40;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
Db 32 DDDDDDD 38

RESULT 36
US-09-864-761-39102
```

```

; Sequence 39102, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39102
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ACO05873.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; US-09-864-761-39102

Query Match      1.5%; Score 7; DB 10; Length 40;
Best Local Similarity 100.0%; Pred. No. 36;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      437 DDDDDDD 443
        |||||
Db      6 DDDDDDD 12
```

```

RESULT 37
; Sequence 35018, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35018
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL049553.14
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.4
; US-09-864-761-35018

Query Match      1.5%; Score 7; DB 10; Length 42;
```



Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 5 DDDDDDD 11

RESULT 38  
US-09-864-761-35629  
; Sequence 35629, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761

PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 35629  
LENGTH: 42  
TYPE: PRT  
ORGANISM: Homo sapiens

FEATURE:  
OTHER INFORMATION: MAP TO AC009492.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.98  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2  
US-09-864-761-35629

Query Match 1.5%; Score 7; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 442  
Db 21 GDDDDDD 27

RESULT 39  
US-09-864-761-39869  
; Sequence 39869, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761

PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 39869  
LENGTH: 51  
TYPE: PRT  
ORGANISM: Homo sapiens

FEATURE:  
OTHER INFORMATION: MAP TO AC004253.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1  
US-09-864-761-39869

Query Match 1.5%; Score 7; DB 10; Length 51;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 1 DDDDDDD 7

## RESULT 40

Sequence 39079, Application US/09864761  
Patent No. US20020048763A1

## GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 39079  
LENGTH: 61  
TYPE: PRT

ORGANISM: Homo sapiens  
FEATURE:

OTHER INFORMATION: MAP TO AC004912.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7  
OTHER INFORMATION: SWISSPROT HIT: Q09322, EVALUUE 3.50e+00  
US-09-864-761-39079

Query Match 1.5%; Score 7; DB 10; Length 61;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 6 DDDDDDD 12

## RESULT 41

Sequence 2, Application US/10028247  
Patent No. US20020150594A1

## GENERAL INFORMATION:

APPLICANT: Goldman, Stanley  
APPLICANT: Lathrop, Stephanie J.  
APPLICANT: Longchamp, Pascal F.  
APPLICANT: Whalen, Robert G.  
APPLICANT: Maxygen, Inc.  
TITLE OF INVENTION: Methods and Compositions for Developing Spore Display  
FILE REFERENCE: 18097A-033520US  
CURRENT APPLICATION NUMBER: US/10/028,247  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: US 60/214,161  
PRIOR FILING DATE: 2000-06-26  
PRIOR APPLICATION NUMBER: US 09/892,208  
PRIOR FILING DATE: 2001-06-26  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 86  
TYPE: PRT  
ORGANISM: Bacillus subtilis  
FEATURE:  
OTHER INFORMATION: Cotc27 including HAI1 epitope region  
US-10-028-247-2

Query Match 1.5%; Score 7; DB 12; Length 86;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 YASLGS 90  
Db 40 YASLGS 46

## RESULT 42

Sequence 4, Application US/09873880  
Patent No. US20020123118A1

## GENERAL INFORMATION:

APPLICANT: Sewalt, Vincent  
APPLICANT: Falco, S. Carl  
APPLICANT: Allen, Stephen M.  
TITLE OF INVENTION: GLYCINE METABOLISM ENZYMES  
FILE REFERENCE: B81192 US CIP  
CURRENT APPLICATION NUMBER: US/09/873,880  
CURRENT FILING DATE: 2001-06-04

;; PRIOR APPLICATION NUMBER: 09/363,321  
;; PRIOR FILING DATE: July 28, 1999  
;; PRIOR APPLICATION NUMBER: 60/094,839  
;; PRIOR FILING DATE: July 31, 1998  
;; NUMBER OF SEQ ID NOS: 42  
;; SOFTWARE: Microsoft Office 97  
;; SEQ ID NO 4  
;; LENGTH: 91  
;; TYPE: PRT  
;; ORGANISM: Zea mays  
;; US-09-873-880-4

Query Match 1.5%; Score 7; DB 10; Length 91;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 SNGEGHG 370  
|||||  
Db 4 SNGEGHG 10

## RESULT 43

US-09-864-761-44209  
; Sequence 44209, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

;; SEQ ID NO 44209  
;; LENGTH: 102  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AC005668.1  
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4  
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4  
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.55  
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1  
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1  
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
;; US-09-864-761-44209

Query Match 1.5%; Score 7; DB 10; Length 102;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
|||||  
Db 35 DDDDDDD 41

## RESULT 44

US-09-864-761-44215  
; Sequence 44215, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29

```
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 42215
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC018801.2.
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.8
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.2
OTHER INFORMATION: EST HUMAN HIT: BF507876.1, EVALUE 7.00e-55
OTHER INFORMATION: SWISSPROT HIT: P34657, EVALUE 3.00e-21
US-09-864-761-42215
```

```
Query Match
Best Local Similarity 1.5%; Score 7; DB 10; Length 110;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 437 DDDDDDD 443
      |||||
Db 103 DDDDDDD 109
```

```
RESULT 45
US-10-101-464A-662
; Sequence 662, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 662
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Plinus radiata
US-10-101-464A-662
```

```
Query Match
Best Local Similarity 1.5%; Score 7; DB 9; Length 154;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 25 GSFGSVY 31
      |||||
Db 27 GSFGSVY 33
```

```
RESULT 46
US-09-864-761-33375
; Sequence 33375, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
```

```
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 33375
LENGTH: 154
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005921.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.9
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
OTHER INFORMATION: SWISSPROT HIT: Q02926, EVALUE 6.30e-01
OTHER INFORMATION: EST_HUMAN HIT: BE741337.1, EVALUE 2.00e-54
OTHER INFORMATION: EST_HUMAN HIT: D78804.1, EVALUE 6.00e-35
US-09-864-761-33375
```

```
Query Match
Best Local Similarity 1.5%; Score 7; DB 10; Length 154;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 437 DDDDDDD 443
      |||||
```



Db 4 DDDDDDD 10

RESULT 47

US-09-864-761-43542  
; Sequence 43542, Application US/09864761  
; Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeomica-X-1  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

CURRENT FILING DATE: US/09/864,761  
CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 43542  
LENGTH: 161  
TYPE: PRT  
ORGANISM: Homo sapiens

FEATURE:  
OTHER INFORMATION: MAP TO AC004816.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.92  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.88  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1  
OTHER INFORMATION: SWISSPROT HIT: P80192, EVALUATE 7.00e-73  
OTHER INFORMATION: EST\_HUMAN HIT: AW452931.1, EVALUATE 4.00e-53  
US-09-864-761-43542

Query Match 1.5%; Score 7; DB 10; Length 161;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 IHRDKS 136  
Db 129 IHRDKS 135

RESULT 48

US-10-101-464A-681

; Sequence 681, Application US/10101464A  
; Publication No. US20030046728A1

GENERAL INFORMATION:

APPLICANT: Strabala, Timothy  
APPLICANT: Nieuwenhuizen, Nicolaas  
APPLICANT: Higgins, Colleen M.

TITLE OF INVENTION: Compositions Isolated from Plant Cells  
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling  
FILE REFERENCE: 11000.1020c2

CURRENT FILING DATE: US/10/101,464A  
CURRENT FILING DATE: 2002-03-18

PRIOR APPLICATION NUMBER: 09/704,302  
PRIOR FILING DATE: 2000-11-01

PRIOR APPLICATION NUMBER: 09/228,986  
PRIOR FILING DATE: 1999-01-12

PRIOR APPLICATION NUMBER: 60/162,866  
PRIOR FILING DATE: 1999-11-01

PRIOR APPLICATION NUMBER: PCT/US00/00724  
PRIOR FILING DATE: 2000-01-11

NUMBER OF SEQ ID NOS: 989  
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 681  
LENGTH: 163  
TYPE: PRT  
ORGANISM: Pinus radiata  
US-10-101-464A-681

Query Match 1.5%; Score 7; DB 9; Length 163;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSFGSVY 31  
Db 99 GSFGSVY 105

RESULT 49

US-09-764-868-816  
; Sequence 816, Application US/09764868  
; Patent No. US20020168711A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT232

CURRENT FILING DATE: US/09/764,868  
CURRENT FILING DATE: 2001-01-17

Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 1510  
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 816  
LENGTH: 170  
TYPE: PRT  
ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: SITE  
LOCATION: (118)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-868-816

Query Match 1.5%; Score 7; DB 9; Length 170;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVI 178

|||||  
Db 94 WMAPEVI 100

RESULT 50  
US-09-764-868-845

; Sequence 845, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 845  
; LENGTH: 170  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-868-845

Query Match 1.5%; Score 7; DB 9; Length 170;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVI 178  
|||||  
Db 96 WMAPEVI 102

RESULT 51  
US-09-764-891-4540

; Sequence 4540, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4540  
; LENGTH: 170  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-891-4540

Query Match 1.5%; Score 7; DB 9; Length 170;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVI 178  
|||||  
Db 96 WMAPEVI 102

RESULT 52  
US-09-764-868-633

; Sequence 633, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 633  
; LENGTH: 174  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-868-633

Query Match 1.5%; Score 7; DB 9; Length 174;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443  
|||||  
Db 165 DDDDDD 171

RESULT 53  
US-09-764-868-833

; Sequence 833, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 833  
; LENGTH: 177  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-868-833

Query Match 1.5%; Score 7; DB 9; Length 177;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVI 178  
|||||  
Db 29 WMAPEVI 35

RESULT 54  
US-09-764-891-5028

; Sequence 5028, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5028  
; LENGTH: 177  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-891-5028

Query Match 1.5%; Score 7; DB 9; Length 177;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVI 178  
|||||  
Db 29 WMAPEVI 35

RESULT 55  
US-10-101-464A-540

```
/ Sequence 540, Application US/10101464A
/ Publication No. US20030046728A1
/ GENERAL INFORMATION:
/ APPLICANT: Strabala, Timothy
/ APPLICANT: Nieuwenhuizen, Nicolaas
/ TITLE OF INVENTION: Compositions Isolated from Plant Cells
/ TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
/ FILE REFERENCE: 11000.1020c2
/ CURRENT APPLICATION NUMBER: US/10/101,464A
/ CURRENT FILING DATE: 2002-03-18
/ PRIOR APPLICATION NUMBER: 09/704,302
/ PRIOR FILING DATE: 2000-11-01
/ PRIOR APPLICATION NUMBER: 09/228,986
/ PRIOR FILING DATE: 1999-01-12
/ PRIOR APPLICATION NUMBER: 60/162,866
/ PRIOR FILING DATE: 1999-11-01
/ PRIOR APPLICATION NUMBER: PCT/US00/00724
/ PRIOR FILING DATE: 2000-01-11
/ NUMBER OF SEQ ID NOS: 989
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 540
/ LENGTH: 188
/ TYPE: PRT
/ ORGANISM: Eucalyptus grandis
/ US-10-101-464A-540
```

```
Query Match 1.5%; Score 7; DB 9; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 129 VIHRLK 135
Db 140 VIHRLK 146
```

```
RESULT 56
US-09-738-626-4510
/ Sequence 4510, Application US/09738626
/ Publication No. US20020197605A1
/ GENERAL INFORMATION:
/ APPLICANT: NAKAGAWA, SATOSHI
/ APPLICANT: MIZOGUCHI, HIROSHI
/ APPLICANT: ANDO, SEIKO
/ APPLICANT: HAYASHI, MIKIRO
/ APPLICANT: OCHIAI, KEIKO
/ APPLICANT: YOKOI, HARUHIKO
/ APPLICANT: TATEISHI, NAOKO
/ APPLICANT: SENOH, AKIHIRO
/ APPLICANT: IKEDA, MASATO
/ APPLICANT: OZAKI, AKIO
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-125
/ CURRENT APPLICATION NUMBER: US/09/738,626
/ CURRENT FILING DATE: 2000-12-18
/ PRIOR APPLICATION NUMBER: JP 99/377484
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: JP 00/159162
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: JP 00/280988
/ PRIOR FILING DATE: 2000-08-03
/ NUMBER OF SEQ ID NOS: 7059
/ SOFTWARE: PatentIn ver. 3.0
/ SEQ ID NO 4510
/ LENGTH: 207
/ TYPE: PRT
/ ORGANISM: Corynebacterium glutamicum
/ US-09-738-626-4510
```

```
Query Match 1.5%; Score 7; DB 9; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 437 DDDDDDD 443
Db 102 DDDDDDD 108
```

```
RESULT 57
US-10-101-464A-702
/ Sequence 702, Application US/10101464A
/ Publication No. US20030046728A1
/ GENERAL INFORMATION:
/ APPLICANT: Strabala, Timothy
/ APPLICANT: Nieuwenhuizen, Nicolaas
/ TITLE OF INVENTION: Compositions Isolated from Plant Cells
/ TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
/ FILE REFERENCE: 11000.1020c2
/ CURRENT APPLICATION NUMBER: US/10/101,464A
/ CURRENT FILING DATE: 2002-03-18
/ PRIOR APPLICATION NUMBER: 09/704,302
/ PRIOR FILING DATE: 2000-11-01
/ PRIOR APPLICATION NUMBER: 09/228,986
/ PRIOR FILING DATE: 1999-01-12
/ PRIOR APPLICATION NUMBER: 60/162,866
/ PRIOR FILING DATE: 1999-11-01
/ PRIOR APPLICATION NUMBER: PCT/US00/00724
/ PRIOR FILING DATE: 2000-01-11
/ NUMBER OF SEQ ID NOS: 989
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 702
/ LENGTH: 238
/ TYPE: PRT
/ ORGANISM: Pinus radiata
/ US-10-101-464A-702
```

```
Query Match 1.5%; Score 7; DB 9; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 129 VIHRLK 135
Db 183 VIHRLK 189
```

```
RESULT 58
US-09-925-301-1424
/ Sequence 1424, Application US/09925301
/ Patent No. US20020052308A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: PA106
/ CURRENT APPLICATION NUMBER: US/09/925,301
/ CURRENT FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05882
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 1694
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1424
/ LENGTH: 244
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (59)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (62)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (221)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

US-09-925-301-1424

Query Match 1.5%; Score 7; DB 10; Length 244;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 328 TPLLLPL 334  
|||||  
Db 15 TPLLLPL 21

RESULT 59

US-09-764-868-1063  
; Sequence 1063, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 1063  
; LENGTH: 250  
; TYPE: PRT

; ORGANISM: Homo sapiens  
US-09-764-868-1063

Query Match 1.5%; Score 7; DB 9; Length 250;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443  
|||||  
Db 241 DDDDDDD 247

RESULT 60

US-09-955-999-77  
; Sequence 77, Application US/09955999  
; Publication No. US20030036505A1  
; GENERAL INFORMATION:

; APPLICANT: Barash et al.  
; TITLE OF INVENTION: Signal Transduction Pathway Component Polynucleotides, Polypeptide  
; TITLE OF INVENTION: Antibodies, and Methods Based Thereon  
; FILE REFERENCE: PT086P1  
; CURRENT APPLICATION NUMBER: US/09/955,999  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 139  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 77  
; LENGTH: 250  
; TYPE: PRT

; ORGANISM: Homo sapiens  
US-09-955-999-77

Query Match 1.5%; Score 7; DB 9; Length 250;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443  
|||||  
Db 241 DDDDDDD 247

RESULT 61

US-09-898-837A-32  
; Sequence 32, Application US/09898837A  
; Publication No. US2003007697A1

; GENERAL INFORMATION:

; APPLICANT: Quinn, Kerry E.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Vernet, Corine  
; APPLICANT: Hermann, John L.  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Taupier Jr., Raymond  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Curagen Corporation  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: Macdougall, John R  
; TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND  
; FILE REFERENCE: 15966-598 CIP  
; CURRENT APPLICATION NUMBER: US/09/898,837A  
; CURRENT FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/165,986  
; PRIOR FILING DATE: 1999-11-17  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/195,637  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347  
; PRIOR FILING DATE: 2000-02-09  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906  
; PRIOR FILING DATE: 2000-07-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 32  
; LENGTH: 256  
; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae  
US-09-898-837A-32

Query Match 1.5%; Score 7; DB 9; Length 256;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 129 VIHRLDK 135  
|||||  
Db 119 VIHRLDK 125

RESULT 62

US-09-949-842-14  
; Sequence 14, Application US/09949842  
; Patent No. US20020164692A1  
; GENERAL INFORMATION:

; APPLICANT: Ni et al.  
; TITLE OF INVENTION: PT047P1  
; FILE REFERENCE: Immune System-Related Polynucleotides, Polypeptides, and Antibodies  
; CURRENT APPLICATION NUMBER: US/09/949,842  
; CURRENT FILING DATE: 2001-09-02  
; PRIOR APPLICATION NUMBER: PCT/US01/07260  
; PRIOR FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: 60/224,367  
; PRIOR FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: 60/187,873  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 257



TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-949-842-14

Query Match  
Best Local Similarity 1.5%; Score 7; DB 9; Length 257;  
Matches 7; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 108 DDDDDDD 114

RESULT 63  
US-09-764-868-821  
Sequence 821, Application US/09764868  
Patent No. US20020168711A1  
GENERAL INFORMATION:

APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PTZ32  
CURRENT APPLICATION NUMBER: US/09/764,868  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 1510  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 821  
LENGTH: 265  
TYPE: PRT  
ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: SITE  
LOCATION: (5)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (11)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (191)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (246)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (251)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (259)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (262)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (264)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-868-821

Query Match  
Best Local Similarity 1.5%; Score 7; DB 9; Length 265;  
Matches 7; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVI 178  
Db 98 WMAPEVI 104

RESULT 64  
US-09-840-704-6  
Sequence 6, Application US/09840704  
Patent No. US20020122801A1  
GENERAL INFORMATION:  
APPLICANT: Dedhar, Shoukat  
APPLICANT: Hannigan, Greg

TITLE OF INVENTION: Interglin-Linked Kinase and its Uses  
FILE REFERENCE: KIN-2CON  
CURRENT APPLICATION NUMBER: US/09/840,704  
CURRENT FILING DATE: 2001-04-23  
PRIOR APPLICATION NUMBER: 09/566,906  
PRIOR FILING DATE: 2000-05-09  
PRIOR APPLICATION NUMBER: US08/752,345  
PRIOR FILING DATE: 1996-11-19  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 271  
TYPE: PRT  
ORGANISM: H. sapiens  
FEATURE:  
NAME/KEY: Other  
LOCATION: (1)...(271)  
US-09-840-704-6

Query Match  
Best Local Similarity 1.5%; Score 7; DB 10; Length 271;  
Matches 7; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 IHRDLKS 136  
Db 122 IHRDLKS 128

RESULT 65  
US-10-274-409-5  
Sequence 5, Application US/10274409  
Publication No. US20030054529A1  
GENERAL INFORMATION:  
APPLICANT: WEBSTER, Marion et al  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
FILE REFERENCE: CL001161DIV  
CURRENT APPLICATION NUMBER: US/10/274,409  
CURRENT FILING DATE: 2002-10-21  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 275  
TYPE: PRT  
ORGANISM: Dictyostelium discoideum  
US-10-274-409-5

Query Match  
Best Local Similarity 1.5%; Score 7; DB 9; Length 275;  
Matches 7; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVI 178  
Db 184 WMAPEVI 190

RESULT 66  
US-10-274-409-6  
Sequence 6, Application US/10274409  
Publication No. US20030054529A1  
GENERAL INFORMATION:  
APPLICANT: WEBSTER, Marion et al  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
FILE REFERENCE: CL001161DIV  
CURRENT APPLICATION NUMBER: US/10/274,409  
CURRENT FILING DATE: 2002-10-21  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 277

```
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-274-409-6

Query Match
Best Local Similarity 1.5%; Score 7; DB 9; Length 277;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 WMAPEVI 178
Db 182 WMAPEVI 188

RESULT 67
US-10-274-409-7
; Sequence 7, Application US/10274409
; Publication No. US20030054529A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001161DIY
; CURRENT APPLICATION NUMBER: US/10/274,409
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-274-409-7

Query Match
Best Local Similarity 1.5%; Score 7; DB 9; Length 277;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 WMAPEVI 178
Db 182 WMAPEVI 188

RESULT 68
US-08-910-386A-14
; Sequence 14, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
```

```
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-910-386A-14

Query Match
Best Local Similarity 1.5%; Score 7; DB 8; Length 280;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 GSFGSVY 31
Db 143 GSFGSVY 149

RESULT 69
US-09-771-161A-106
; Sequence 106, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 106
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-106

Query Match
Best Local Similarity 1.5%; Score 7; DB 10; Length 284;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 130 IHRDLKS 136
Db 276 IHRDLKS 282

RESULT 70
US-10-060-019-22
; Sequence 22, Application US/10060019
; Publication No. US20030003564A1
; GENERAL INFORMATION:
; APPLICANT: Tyers, Mike
; APPLICANT: Willems, Andrew
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN
; TITLE OF INVENTION: DEPENDENT PROTEOLYSIS
; FILE REFERENCE: 11757.10USU1
; CURRENT APPLICATION NUMBER: US/10/060,019
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US/09/177,165
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/092,443
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/063,254
; PRIOR FILING DATE: 1997-10-24
```

NUMBER OF SEQ ID NOS: 50  
SOFTWARE: Patentn Ver. 2.1  
SEQ ID NO 22  
LENGTH: 295  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-10-060-019-22

Query Match 1.5%; Score 7; DB 9; Length 295;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
DB 255 DDDDDDD 261

RESULT 71  
US-09-991-211-5  
Sequence 5, Application US/09991211  
Patent No. US20020156585A1

GENERAL INFORMATION:  
APPLICANT: Pharmacia & Upjohn  
TITLE OF INVENTION: CRYSTALLIZATION AND STRUCTURE DETERMINATION OF  
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS  
TITLE OF INVENTION: UDP-N-ACETYLENOLPYRUVYLGLUCOSAMINE REDUCTASE (S. aureus  
TITLE OF INVENTION: MuB)  
FILE REFERENCE: 268.6241.0101  
CURRENT APPLICATION NUMBER: US/09/991,211  
CURRENT FILING DATE: 2001-11-21  
PRIOR APPLICATION NUMBER: 09/632,947  
PRIOR FILING DATE: 2000-08-04  
PRIOR APPLICATION NUMBER: 60/147,164  
PRIOR FILING DATE: 1999-08-04  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: Patentn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 303  
TYPE: PRT  
ORGANISM: Bacillus subtilis  
US-09-991-211-5

Query Match 1.5%; Score 7; DB 9; Length 303;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 QELKERE 312  
DB 6 QELKERE 12

RESULT 72  
US-09-866-582-27  
Sequence 27, Application US/09866582  
Patent No. US20020127620A1  
GENERAL INFORMATION:  
APPLICANT: Wiltman, George B.  
APPLICANT: Pazour, Gregory J.  
APPLICANT: Rosenbaum, Joel L.  
APPLICANT: Cole, Douglas G.  
TITLE OF INVENTION: INTRAFILLAGELLAR TRANSPORT  
FILE REFERENCE: 07917-145001  
CURRENT APPLICATION NUMBER: US/09/866,582  
CURRENT FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: US 60/206,923  
PRIOR FILING DATE: 2000-05-24  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 27  
LENGTH: 304  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-866-582-27

Query Match 1.5%; Score 7; DB 10; Length 304;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
DB 34 DDDDDDD 40

RESULT 73  
US-09-746-801A-61  
Sequence 61, Application US/09746801A  
Patent No. US20020083494A1

GENERAL INFORMATION:  
APPLICANT: Wagner, et al.  
TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM  
FILE REFERENCE: 1505-54357  
CURRENT APPLICATION NUMBER: US/09/746,801A  
CURRENT FILING DATE: 2000-12-20  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: Patentn version 3.1  
SEQ ID NO 61  
LENGTH: 324  
TYPE: PRT  
ORGANISM: Brassica sp.  
NAME/KEY: misc feature  
LOCATION: (188)..(188)  
OTHER INFORMATION: Xaa = uncertain amino acid residue  
NAME/KEY: misc feature  
LOCATION: (194)..(194)  
OTHER INFORMATION: Xaa = uncertain amino acid residue  
NAME/KEY: misc feature  
LOCATION: (196)..(196)  
OTHER INFORMATION: Xaa = uncertain amino acid residue  
NAME/KEY: misc feature  
LOCATION: (214)..(214)  
OTHER INFORMATION: Xaa = uncertain amino acid residue  
US-09-746-801A-61

Query Match 1.5%; Score 7; DB 10; Length 324;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
DB 162 DDDDDDD 168

RESULT 74  
US-09-815-242-5783  
Sequence 5783, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727

Search completed: May 1, 2003, 21:01:50  
Job time: 54 secs

```
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5783
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5783
```

```
Query Match 1.5%; Score 7; DB 10; Length 325;
Best Local Similarity 100.0%; Pred.No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

57 LSVLSHR 63  
Db 252 LSVLSHR 258

```
RESULT 75
US-09-815-242-12693
; Sequence 12693, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyckind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.01A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12693
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12693
```

```
Query Match 1.5%; Score 7; DB 10; Length 325;
Best Local Similarity 100.0%; Pred.No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

57 LSVLSHR 63  
Db 252 LSVLSHR 258





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:52:23 ; Search time 19 Seconds  
(without alignments)  
704.602 Million cell updates/sec

Title: US-09-757-982-5  
Perfect score: 455  
Sequence: 1 MSSLGASFVQIKFDLQFFE.....GDDDDDDGEEEDNDMDNSE 455

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

rd size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep: \*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep: \*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep: \*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                          |
|------------|-------|-------------|--------|-------|--------------------------------------|
| 1          | 455   | 100.0       | 455    | 3     | US-09-221-235-5 Sequence 5, Appli    |
| 2          | 455   | 100.0       | 455    | 3     | US-09-221-928-5 Sequence 5, Appli    |
| 3          | 455   | 100.0       | 455    | 4     | US-09-221-527-5 Sequence 5, Appli    |
| 4          | 455   | 100.0       | 455    | 4     | US-09-221-236-5 Sequence 5, Appli    |
| 5          | 455   | 100.0       | 455    | 4     | US-09-221-416-5 Sequence 5, Appli    |
| 6          | 455   | 100.0       | 455    | 4     | US-09-221-245-5 Sequence 5, Appli    |
| 7          | 455   | 100.0       | 455    | 4     | US-09-163-115-5 Sequence 5, Appli    |
| 8          | 455   | 100.0       | 455    | 4     | US-09-221-528-5 Sequence 5, Appli    |
| 9          | 455   | 100.0       | 455    | 4     | US-09-593-553-5 Sequence 5, Appli    |
| 10         | 455   | 100.0       | 455    | 4     | US-09-221-237-5 Sequence 5, Appli    |
| 11         | 455   | 100.0       | 455    | 1     | US-08-375-300-4 Sequence 5, Appli    |
| 12         | 455   | 100.0       | 455    | 3     | US-09-177-431-4 Sequence 4, Appli    |
| 13         | 455   | 100.0       | 455    | 5     | PCT-US95-16930-4 Sequence 4, Appli   |
| 14         | 455   | 100.0       | 455    | 1     | US-08-375-300-2 Sequence 2, Appli    |
| 15         | 455   | 100.0       | 455    | 3     | US-09-177-431-2 Sequence 2, Appli    |
| 16         | 455   | 100.0       | 455    | 5     | PCT-US95-16930-2 Sequence 2, Appli   |
| 17         | 455   | 100.0       | 455    | 4     | US-08-158-735A-15 Sequence 15, Appli |
| 18         | 455   | 100.0       | 455    | 4     | US-09-382-256-12 Sequence 12, Appli  |
| 19         | 455   | 100.0       | 455    | 4     | US-09-395-115-12 Sequence 12, Appli  |
| 20         | 455   | 100.0       | 455    | 4     | US-08-436-265-12 Sequence 12, Appli  |
| 21         | 455   | 100.0       | 455    | 4     | US-09-679-187-12 Sequence 12, Appli  |
| 22         | 455   | 100.0       | 455    | 1     | US-08-149-105-16 Sequence 16, Appli  |
| 23         | 455   | 100.0       | 455    | 1     | US-08-317-847-16 Sequence 16, Appli  |
| 24         | 455   | 100.0       | 455    | 2     | US-08-446-855A-2 Sequence 2, Appli   |
| 25         | 455   | 100.0       | 455    | 4     | US-09-150-741-2 Sequence 2, Appli    |
| 26         | 455   | 100.0       | 455    | 1     | US-08-240-712-35 Sequence 35, Appli  |
| 27         | 455   | 100.0       | 455    | 1     | US-08-443-890-35 Sequence 35, Appli  |

|     |     |      |   |                     |                    |
|-----|-----|------|---|---------------------|--------------------|
| 28  | 1.8 | 43   | 4 | US-09-303-814-1     | Sequence 1, Appli  |
| 29  | 1.8 | 158  | 2 | US-08-618-911-4     | Sequence 4, Appli  |
| 30  | 1.8 | 158  | 2 | US-08-618-911-6     | Sequence 6, Appli  |
| 31  | 1.8 | 158  | 3 | US-08-938-675A-2    | Sequence 2, Appli  |
| 32  | 1.8 | 179  | 2 | US-08-752-132-2     | Sequence 2, Appli  |
| 33  | 1.8 | 192  | 3 | US-09-382-080-1     | Sequence 1, Appli  |
| 34  | 1.8 | 192  | 3 | US-08-859-937-1     | Sequence 26, Appli |
| 35  | 1.8 | 226  | 1 | US-08-431-080-26    | Sequence 26, Appli |
| 36  | 1.8 | 226  | 2 | US-08-938-534-26    | Sequence 26, Appli |
| 37  | 1.8 | 226  | 4 | US-09-345-294-26    | Sequence 26, Appli |
| 38  | 1.8 | 251  | 2 | US-08-469-537A-80   | Sequence 80, Appli |
| 39  | 1.8 | 251  | 2 | US-08-469-537A-81   | Sequence 81, Appli |
| 40  | 1.8 | 270  | 2 | US-08-852-743-5     | Sequence 5, Appli  |
| 41  | 1.8 | 270  | 3 | US-09-185-370-5     | Sequence 5, Appli  |
| 42  | 1.8 | 317  | 4 | US-09-485-077A-17   | Sequence 17, Appli |
| 43  | 1.8 | 416  | 2 | US-09-211-930-3     | Sequence 3, Appli  |
| 44  | 1.8 | 416  | 2 | US-09-211-930-11    | Sequence 11, Appli |
| 45  | 1.8 | 416  | 3 | US-09-340-993-3     | Sequence 3, Appli  |
| 46  | 1.8 | 416  | 3 | US-09-340-993-11    | Sequence 11, Appli |
| 47  | 1.8 | 416  | 4 | US-09-468-442-3     | Sequence 3, Appli  |
| 48  | 1.8 | 416  | 4 | US-09-468-442-11    | Sequence 11, Appli |
| 49  | 1.8 | 487  | 2 | US-08-712-709-8     | Sequence 8, Appli  |
| 50  | 1.8 | 487  | 3 | US-09-111-444-8     | Sequence 8, Appli  |
| 51  | 1.8 | 487  | 4 | US-09-541-228-8     | Sequence 8, Appli  |
| 52  | 1.8 | 579  | 4 | US-09-529-279-4     | Sequence 4, Appli  |
| 53  | 1.8 | 590  | 4 | US-09-529-279-15    | Sequence 15, Appli |
| 54  | 1.8 | 943  | 2 | US-08-469-537A-107  | Sequence 107, App  |
| 55  | 1.8 | 1085 | 1 | US-08-431-080-28    | Sequence 28, Appli |
| 56  | 1.8 | 1085 | 2 | US-08-938-534-28    | Sequence 28, Appli |
| 57  | 1.8 | 1085 | 4 | US-09-345-294-28    | Sequence 28, Appli |
| 58  | 1.5 | 11   | 2 | US-08-482-228-215   | Sequence 215, App  |
| 59  | 1.5 | 11   | 3 | US-08-199-776-23    | Sequence 23, Appli |
| 60  | 1.5 | 12   | 1 | US-08-199-776-23    | Sequence 23, Appli |
| 61  | 1.5 | 12   | 3 | US-08-663-731-23    | Sequence 23, Appli |
| 62  | 1.5 | 12   | 4 | US-09-293-238B-23   | Sequence 23, Appli |
| 63  | 1.5 | 12   | 5 | PCT-US95-02044-23   | Sequence 23, Appli |
| 64  | 1.5 | 13   | 4 | US-09-625-570-4     | Sequence 4, Appli  |
| 65  | 1.5 | 14   | 4 | US-09-039-780A-80   | Sequence 80, Appli |
| 66  | 1.5 | 30   | 4 | US-09-039-780A-87   | Sequence 87, Appli |
| 67  | 1.5 | 30   | 4 | US-09-376-113-3     | Sequence 3, Appli  |
| 68  | 1.5 | 50   | 4 | US-08-290-736C-9    | Sequence 9, Appli  |
| 69  | 1.5 | 73   | 4 | US-08-117-952-757   | Sequence 757, App  |
| 70  | 1.5 | 76   | 2 | US-08-376-113-2     | Sequence 2, Appli  |
| 71  | 1.5 | 86   | 1 | US-08-370-225-20    | Sequence 20, Appli |
| 72  | 1.5 | 86   | 1 | US-08-461-859-20    | Sequence 20, Appli |
| 73  | 1.5 | 86   | 5 | PCT-US93-10069-20   | Sequence 20, Appli |
| 74  | 1.5 | 118  | 1 | US-08-072-574-4     | Sequence 4, Appli  |
| 75  | 1.5 | 118  | 1 | US-08-486-270-4     | Sequence 4, Appli  |
| 76  | 1.5 | 118  | 3 | US-08-367-264-4     | Sequence 4, Appli  |
| 77  | 1.5 | 118  | 4 | US-09-153-757-4     | Sequence 4, Appli  |
| 78  | 1.5 | 126  | 4 | US-08-483-533-31    | Sequence 31, Appli |
| 79  | 1.5 | 126  | 4 | US-09-283-471A-31   | Sequence 31, Appli |
| 80  | 1.5 | 159  | 1 | US-08-463-090B-12   | Sequence 12, Appli |
| 81  | 1.5 | 168  | 4 | US-09-376-113-5     | Sequence 5, Appli  |
| 82  | 1.5 | 169  | 4 | US-08-483-533-28    | Sequence 28, Appli |
| 83  | 1.5 | 172  | 1 | US-08-318-947A-11   | Sequence 11, Appli |
| 84  | 1.5 | 172  | 2 | US-08-795-303-11    | Sequence 11, Appli |
| 85  | 1.5 | 172  | 3 | US-09-358-580-14    | Sequence 14, Appli |
| 86  | 1.5 | 175  | 4 | US-09-376-113-7     | Sequence 7, Appli  |
| 87  | 1.5 | 179  | 4 | US-08-483-533-33    | Sequence 33, Appli |
| 88  | 1.5 | 183  | 1 | US-08-176-620A-8    | Sequence 8, Appli  |
| 89  | 1.5 | 183  | 2 | US-08-461-985-8     | Sequence 8, Appli  |
| 90  | 1.5 | 183  | 4 | US-08-932-012C-16   | Sequence 16, Appli |
| 91  | 1.5 | 183  | 4 | US-08-888-818C-16   | Sequence 16, Appli |
| 92  | 1.5 | 183  | 4 | US-08-483-533-42    | Sequence 42, Appli |
| 93  | 1.5 | 245  | 4 | US-09-283-471A-42   | Sequence 42, Appli |
| 94  | 1.5 | 246  | 4 | US-09-134-001C-5070 | Sequence 5070, Ap  |
| 95  | 1.5 | 252  | 4 | US-08-483-533-43    | Sequence 43, Appli |
| 96  | 1.5 |      |   |                     |                    |
| 97  | 1.5 |      |   |                     |                    |
| 98  | 1.5 |      |   |                     |                    |
| 99  | 1.5 |      |   |                     |                    |
| 100 | 1.5 |      |   |                     |                    |

ALIGNMENTS

```
RESULT 1
; US-09-221-235-5
; Sequence 5, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-221-235-5

Query Match          100.0%; Score 455; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSODKEVAVKKLKEKEAELISVL 60
    |||
Db 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSODKEVAVKKLKEKEAELISVL 60

QY 61 SHRNIQFYGVILLEPPNYGI VTEYASLSGLYDINSNRSEEMDMHMTWATDVAKGMHY 120
    |||
Db 61 SHRNIQFYGVILLEPPNYGI VTEYASLSGLYDINSNRSEEMDMHMTWATDVAKGMHY 120

QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180
    |||
Db 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180

QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLQVAVMLVVEKNERLTI PSSCPRSFAELH 240
    |||
Db 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLQVAVMLVVEKNERLTI PSSCPRSFAELH 240

QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLDPKCN SFLHKAEMWRCIEATLERLKLKLERD 300
    |||
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLDPKCN SFLHKAEMWRCIEATLERLKLKLERD 300

QY 301 LSFKEQELKERERLKMWEQKLTQSNTPILLPLAARMSEESYFESKTEESNSAEMSCQI 360
    |||
Db 301 LSFKEQELKERERLKMWEQKLTQSNTPILLPLAARMSEESYFESKTEESNSAEMSCQI 360

QY 361 TATSNGEHGMNPSLOAMLMGF GDI FSNMKA GAVMHS GMOINMOAKONSSKTTSKRGRK 420
    |||
Db 361 TATSNGEHGMNPSLOAMLMGF GDI FSNMKA GAVMHS GMOINMOAKONSSKTTSKRGRK 420

QY 421 KVMALGFSDFDLSEGD DDDDDGEEEDNDMDNSE 455
    |||
Db 421 KVMALGFSDFDLSEGD DDDDDGEEEDNDMDNSE 455

RESULT 2
; US-09-221-928-5
; Sequence 5, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
```

```
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-221-928-5

Query Match          100.0%; Score 455; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSODKEVAVKKLKEKEAELISVL 60
    |||
Db 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSODKEVAVKKLKEKEAELISVL 60

QY 61 SHRNIQFYGVILLEPPNYGI VTEYASLSGLYDINSNRSEEMDMHMTWATDVAKGMHY 120
    |||
Db 61 SHRNIQFYGVILLEPPNYGI VTEYASLSGLYDINSNRSEEMDMHMTWATDVAKGMHY 120

QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180
    |||
Db 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180

QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLQVAVMLVVEKNERLTI PSSCPRSFAELH 240
    |||
Db 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLQVAVMLVVEKNERLTI PSSCPRSFAELH 240

QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLDPKCN SFLHKAEMWRCIEATLERLKLKLERD 300
    |||
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLDPKCN SFLHKAEMWRCIEATLERLKLKLERD 300

QY 301 LSFKEQELKERERLKMWEQKLTQSNTPILLPLAARMSEESYFESKTEESNSAEMSCQI 360
    |||
Db 301 LSFKEQELKERERLKMWEQKLTQSNTPILLPLAARMSEESYFESKTEESNSAEMSCQI 360

QY 361 TATSNGEHGMNPSLOAMLMGF GDI FSNMKA GAVMHS GMOINMOAKONSSKTTSKRGRK 420
    |||
Db 361 TATSNGEHGMNPSLOAMLMGF GDI FSNMKA GAVMHS GMOINMOAKONSSKTTSKRGRK 420

QY 421 KVMALGFSDFDLSEGD DDDDDGEEEDNDMDNSE 455
    |||
Db 421 KVMALGFSDFDLSEGD DDDDDGEEEDNDMDNSE 455

RESULT 3
; US-09-221-527-5
; Sequence 5, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-221-527-5

Query Match          100.0%; Score 455; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSODKEVAVKKLKEKEAELISVL 60
    |||
Db 1 MSSLGASFVQIKFDDLOFFENC GGSGSVYRAKWSODKEVAVKKLKEKEAELISVL 60
```

```
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMDHIMTWATDVAKGMHY 120
DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMDHIMTWATDVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180
DB 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180
QY 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVMLVVEKNERLTISSCPRSFAELH 240
DB 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVMLVVEKNERLTISSCPRSFAELH 240
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATLERLKLKLERD 300
DB 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATLERLKLKLERD 300
QY 301 LSFKEQLKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
DB 301 LSFKEQLKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
QY 361 TATSNGEHGMNPSLOAMLMFGDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420
DB 361 TATSNGEHGMNPSLOAMLMFGDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420
QY 421 KVNMLGFSDFDLSEGDGDDDDGEEEDNDMDNSE 455
DB 421 KVNMLGFSDFDLSEGDGDDDDGEEEDNDMDNSE 455
```

## RESULT 4

```
US-09-221-236-5
; Sequence 5, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PR1
; ORGANISM: Homo sapiens
-09-221-236-5
```

Query Match 100.0%; Score 455; DB 4; Length 455;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWIISQDKEVAVKLLKIEKEAIIISVL 60
DB 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWIISQDKEVAVKLLKIEKEAIIISVL 60
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMDHIMTWATDVAKGMHY 120
DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMDHIMTWATDVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180
DB 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180
QY 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVMLVVEKNERLTISSCPRSFAELH 240
DB 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVMLVVEKNERLTISSCPRSFAELH 240
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATLERLKLKLERD 300
DB 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATLERLKLKLERD 300
```

```
QY 301 LSFKEQLKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
DB 301 LSFKEQLKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
QY 361 TATSNGEHGMNPSLOAMLMFGDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420
DB 361 TATSNGEHGMNPSLOAMLMFGDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420
QY 421 KVNMLGFSDFDLSEGDGDDDDGEEEDNDMDNSE 455
DB 421 KVNMLGFSDFDLSEGDGDDDDGEEEDNDMDNSE 455
```

## RESULT 5

```
US-09-221-416-5
; Sequence 5, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-221-416-5
```

Query Match 100.0%; Score 455; DB 4; Length 455;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWIISQDKEVAVKLLKIEKEAIIISVL 60
DB 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWIISQDKEVAVKLLKIEKEAIIISVL 60
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMDHIMTWATDVAKGMHY 120
DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSNRSEEMDMDHIMTWATDVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180
DB 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCICDFGASRFHNHTTHMSLVGTFPMMAPEVIQS 180
QY 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVMLVVEKNERLTISSCPRSFAELH 240
DB 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAVMLVVEKNERLTISSCPRSFAELH 240
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATLERLKLKLERD 300
DB 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFLHNKAERCEIATLERLKLKLERD 300
QY 301 LSFKEQLKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
DB 301 LSFKEQLKERERRLKMEQKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360
QY 361 TATSNGEHGMNPSLOAMLMFGDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420
DB 361 TATSNGEHGMNPSLOAMLMFGDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420
QY 421 KVNMLGFSDFDLSEGDGDDDDGEEEDNDMDNSE 455
DB 421 KVNMLGFSDFDLSEGDGDDDDGEEEDNDMDNSE 455
```

## RESULT 6

US-09-221-245-5



```
; Sequence 5, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-221-245-5
```

```
Query Match 100.0%; Score 455; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MSSLGASFVQIKFDDLOFFENCGGSGSVYRAKWSQDKEVAVKLLKIEKAEILSVL 60
Db 1 MSSLGASFVQIKFDDLOFFENCGGSGSVYRAKWSQDKEVAVKLLKIEKAEILSVL 60
QY 61 SHRNIQFYGVILEPPNYGIATEYASLSGLYDINSNRSEEMDMHIMTWATVAKGMHY 120
Db 61 SHRNIQFYGVILEPPNYGIATEYASLSGLYDINSNRSEEMDMHIMTWATVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180
Db 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180
QY 181 LPVSETCDITYSYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPRSFAELH 240
Db 181 LPVSETCDITYSYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPRSFAELH 240
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFHLNKAEMRCEIATLERLKLERD 300
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFHLNKAEMRCEIATLERLKLERD 300
QY 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQI 360
Db 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQI 360
QY 361 TATSNGEHGMNPSLOAMLMGFDFSMNKAGAVMHSQMOMQAKONSSKTSKRGRK 420
Db 361 TATSNGEHGMNPSLOAMLMGFDFSMNKAGAVMHSQMOMQAKONSSKTSKRGRK 420
QY 421 KVMALGFSDFDLSGDDDDDDGEEEDNDMDNSE 455
Db 421 KVMALGFSDFDLSGDDDDDDGEEEDNDMDNSE 455
```

```
RESULT 7
US-09-163-115-5
; Sequence 5, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-163-115-5
```

```
Query Match 100.0%; Score 455; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MSSLGASFVQIKFDDLOFFENCGGSGSVYRAKWSQDKEVAVKLLKIEKAEILSVL 60
Db 1 MSSLGASFVQIKFDDLOFFENCGGSGSVYRAKWSQDKEVAVKLLKIEKAEILSVL 60
QY 61 SHRNIQFYGVILEPPNYGIATEYASLSGLYDINSNRSEEMDMHIMTWATVAKGMHY 120
Db 61 SHRNIQFYGVILEPPNYGIATEYASLSGLYDINSNRSEEMDMHIMTWATVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180
Db 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180
QY 181 LPVSETCDITYSYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPRSFAELH 240
Db 181 LPVSETCDITYSYGVVLWEMLTREVPFKGLEGLQVAMLVVEKNERLTI PSSCPRSFAELH 240
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFHLNKAEMRCEIATLERLKLERD 300
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKCNFHLNKAEMRCEIATLERLKLERD 300
QY 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQI 360
Db 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQI 360
QY 361 TATSNGEHGMNPSLOAMLMGFDFSMNKAGAVMHSQMOMQAKONSSKTSKRGRK 420
Db 361 TATSNGEHGMNPSLOAMLMGFDFSMNKAGAVMHSQMOMQAKONSSKTSKRGRK 420
QY 421 KVMALGFSDFDLSGDDDDDDGEEEDNDMDNSE 455
Db 421 KVMALGFSDFDLSGDDDDDDGEEEDNDMDNSE 455
```

```
RESULT 8
US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-221-528-5
```

```
Query Match 100.0%; Score 455; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENCGGSGSVYRAKWSQDKEVAVKLLKIEKAEILSVL 60
Db 1 MSSLGASFVQIKFDDLOFFENCGGSGSVYRAKWSQDKEVAVKLLKIEKAEILSVL 60
QY 61 SHRNIQFYGVILEPPNYGIATEYASLSGLYDINSNRSEEMDMHIMTWATVAKGMHY 120
Db 61 SHRNIQFYGVILEPPNYGIATEYASLSGLYDINSNRSEEMDMHIMTWATVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180
Db 121 LHMEAPVKVIHRDLKSRNVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMAPEVIQS 180
```

QY 181 LPVSETCDTYSYGVLMWMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240  
Db 181 LPVSETCDTYSYGVLMWMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKLKLERD 300  
QY 301 LSFKEQELKERERRLKMEQKLTQOSNTPLPLAARMSEESYFESKTEESNSAEMSCQI 360  
Db 301 LSFKEQELKERERRLKMEQKLTQOSNTPLPLAARMSEESYFESKTEESNSAEMSCQI 360  
QY 361 TATSNGEHGMNPSLQAMLMGFGDIFSMNKAGAVMHSGMOINMOAKONSSKTSKRGRK 420  
Db 361 TATSNGEHGMNPSLQAMLMGFGDIFSMNKAGAVMHSGMOINMOAKONSSKTSKRGRK 420  
QY 421 KVNMLGFSDFDLSGGDDDDDDGEEEDNDMDNSE 455  
Db 421 KVNMLGFSDFDLSGGDDDDDDGEEEDNDMDNSE 455

## RESULT 9

US-09-593-553-5  
; Sequence 5, Application US/09593553  
; Patent No. 6200770

## GENERAL INFORMATION:

; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/593,553  
; CURRENT FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: 09/163,115  
; PRIOR FILING DATE: 1998-09-28  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-593-553-5

Query Match 100.0%; Score 455; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWSIQDKEVAVKLLIKEAEILSVL 60  
Db 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWSIQDKEVAVKLLIKEAEILSVL 60  
QY 61 SHRNIQFYGVILEPPNYGIVTEYASIGSLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
Db 61 SHRNIQFYGVILEPPNYGIVTEYASIGSLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
QY 181 LPVSETCDTYSYGVLMWMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240  
Db 181 LPVSETCDTYSYGVLMWMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKLKLERD 300  
QY 301 LSFKEQELKERERRLKMEQKLTQOSNTPLPLAARMSEESYFESKTEESNSAEMSCQI 360  
Db 301 LSFKEQELKERERRLKMEQKLTQOSNTPLPLAARMSEESYFESKTEESNSAEMSCQI 360  
QY 361 TATSNGEHGMNPSLQAMLMGFGDIFSMNKAGAVMHSGMOINMOAKONSSKTSKRGRK 420  
Db 361 TATSNGEHGMNPSLQAMLMGFGDIFSMNKAGAVMHSGMOINMOAKONSSKTSKRGRK 420

QY 421 KVNMLGFSDFDLSGGDDDDDDGEEEDNDMDNSE 455  
Db 421 KVNMLGFSDFDLSGGDDDDDDGEEEDNDMDNSE 455

## RESULT 10

US-09-221-237-5

; Sequence 5, Application US/09221237

; Patent No. 6214597

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050

; CURRENT APPLICATION NUMBER: US/09/221,237

; EARLIER APPLICATION NUMBER: 09/163,115

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 455

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-221-237-5

Query Match 100.0%; Score 455; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWSIQDKEVAVKLLIKEAEILSVL 60  
Db 1 MSSLGASFVQIKFDDLOFFENCGGSGFSVYRAKWSIQDKEVAVKLLIKEAEILSVL 60  
QY 61 SHRNIQFYGVILEPPNYGIVTEYASIGSLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
Db 61 SHRNIQFYGVILEPPNYGIVTEYASIGSLYDINSNRSEEMDMHMTWATDVAKGMHY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTHMSLVGTFPMAPEVIQS 180  
QY 181 LPVSETCDTYSYGVLMWMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240  
Db 181 LPVSETCDTYSYGVLMWMLTREVPFKGLEGLQVAMLVVEKNERLTIPSSCPSPFAELH 240  
QY 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEMRCEIATLERLKLKLERD 300  
QY 301 LSFKEQELKERERRLKMEQKLTQOSNTPLPLAARMSEESYFESKTEESNSAEMSCQI 360  
Db 301 LSFKEQELKERERRLKMEQKLTQOSNTPLPLAARMSEESYFESKTEESNSAEMSCQI 360  
QY 361 TATSNGEHGMNPSLQAMLMGFGDIFSMNKAGAVMHSGMOINMOAKONSSKTSKRGRK 420  
Db 361 TATSNGEHGMNPSLQAMLMGFGDIFSMNKAGAVMHSGMOINMOAKONSSKTSKRGRK 420  
QY 421 KVNMLGFSDFDLSGGDDDDDDGEEEDNDMDNSE 455  
Db 421 KVNMLGFSDFDLSGGDDDDDDGEEEDNDMDNSE 455

## RESULT 11

US-08-375-300-4

; Sequence 4, Application US/08375300

; Patent No. 5679566

; GENERAL INFORMATION:

; APPLICANT: Feng, He

; TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN

; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street Suite 3100  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/375,300  
FILING DATE: 20-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. P.  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04020/046001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)542-5070  
TELEFAX: (617)542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 764 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-375-300-4

Query Match 2.2%; Score 10; DB 1; Length 764;  
Best Local Similarity 100.0%; Pred. No. 0.64;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGEE 446  
|||||  
Db 577 DDDDDDDGEE 586

RESULT 12  
US-09-177-431-4  
Sequence 4, Application US/09177431  
Patent No. 6071700  
GENERAL INFORMATION:  
APPLICANT: He, Feng  
APPLICANT: Jacobson, Allan S.  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE  
TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/177,431  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/955,472  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 07917/050001  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070  
TELEFAX: 617/542-9806  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 764 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-177-431-4

Query Match 2.2%; Score 10; DB 3; Length 764;  
Best Local Similarity 100.0%; Pred. No. 0.64;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGEE 446  
|||||  
Db 577 DDDDDDDGEE 586

RESULT 13  
PCT-US95-16930-4  
Sequence 4, Application PC/TUS9516930  
GENERAL INFORMATION:  
APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE  
TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF  
TITLE OF INVENTION: NONSENSE-MEDIATED mRNA DECAY  
TITLE OF INVENTION: FUNCTION  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street Suite 3100  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/16930  
FILING DATE: 27-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375,300  
FILING DATE: 20-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04020/046001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)542-5070  
TELEFAX: (617)542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 764 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-16930-4

Query Match 2.2%; Score 10; DB 5; Length 764;  
Best Local Similarity 100.0%; Pred. No. 0.64;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGEE 446  
|||||  
Db 577 DDDDDDDGEE 586

RESULT 14  
US-08-375-300-2  
Sequence 2, Application US/08375300  
Patent No. 5679566  
GENERAL INFORMATION:  
APPLICANT: Feng, He  
APPLICANT: Jacobson, Allan S.  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN  
TITLE OF INVENTION: THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street Suite 3100  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/375,300  
FILING DATE: 20-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. P.  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04020/046001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)542-5070  
TELEFAX: (617)542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1089 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-375-300-2

Query Match 2.2%; Score 10; DB 1; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 0.86;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

437 DDDDDDDGEE 446  
|||||  
902 DDDDDDDGEE 911

RESULT 15  
US-09-177-431-2  
Sequence 2, Application US/09177431  
Patent No. 6071700  
GENERAL INFORMATION:  
APPLICANT: He, Feng  
APPLICANT: Jacobson, Allan S.  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE  
TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95

SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/177,431  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/955,472  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 07917/050001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-9806  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1089 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-177-431-2

Query Match 2.2%; Score 10; DB 3; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 0.86;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446  
|||||  
Db 902 DDDDDDDGEE 911

RESULT 16  
PCT-US95-16930-2  
Sequence 2, Application PC/TUS9516930  
GENERAL INFORMATION:  
APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE  
TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF  
TITLE OF INVENTION: NONSENSE-MEDIATED mRNA DECAY  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street Suite 3100  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/16930  
FILING DATE: 27-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375,300  
FILING DATE: 20-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04020/046001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)542-5070  
TELEFAX: (617)542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1089 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant



TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-16930-2

Query Match 2.2%; Score 10; DB 5; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 0.86;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446  
| | | | |  
Db 902 DDDDDDDGEE 911

RESULT 17  
US-08-158-735A-15  
Sequence 15, Application US/08158735A  
Patent No. 6248554

GENERAL INFORMATION:

APPLICANT: COOK, JONATHAN S.

APPLICANT: CORREA, PAUL E.

APPLICANT: KOENIG, BETH B.

APPLICANT: ROSENBAUM, JAN S.

APPLICANT: TING, JERRY

TITLE OF INVENTION: DNA SEQUENCE CODING FOR A BMP RECEPTOR

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: THE PROCTER & GAMBLE COMPANY

STREET: 11810 EAST MIAMI RIVER ROAD

CITY: ROSS

STATE: OH

COUNTRY: USA

ZIP: 45061

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/158,735A

FILING DATE: 24-NOV-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CORSTANDE, BRAHM J.

REGISTRATION NUMBER: 34,804

REFERENCE/DOCKET NUMBER: 5088

TELECOMMUNICATION INFORMATION:

TELEPHONE: (513) 627-2858

TELEFAX: (513) 627-0260

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 303 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-158-735A-15

Query Match 2.0%; Score 9; DB 4; Length 303;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
| | | | |  
Db 128 HRDLKSRNV 136

RESULT 18  
US-09-382-256-12  
Sequence 12, Application US/09382256A  
Patent No. 6207814

GENERAL INFORMATION:  
APPLICANT: MIYAZONO, Kohei  
TEN DIJKE, Peter

FRANZEN, Petra  
YAMASHITA, Hidetoshi  
HELDIN, Carl-Henrik  
TITLE OF INVENTION: ACTIVIN RECEPTOR LIKE KINASES, PROTEINS,  
HAVING SERINE THREONINE KINASE DOMAINS,  
AND THEIR USE

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fulbright & Jaworski L.L.P.

STREET: 666 Fifth Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10103

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/382,256A

FILING DATE: 24-Aug-1999

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB93/02367

FILING DATE: No. 6207814ember 17, 1993

APPLICATION NUMBER: GB 9224057.1

FILING DATE: No. 6207814ember 17, 1992

APPLICATION NUMBER: GB 9304677.9

FILING DATE: March 8, 1993

APPLICATION NUMBER: GB 9304680.3

FILING DATE: March 8, 1993

APPLICATION NUMBER: 9311047.6

FILING DATE: May 28, 1993

APPLICATION NUMBER: 9313763.6

FILING DATE: July 2, 1993

APPLICATION NUMBER: 9316099.2

FILING DATE: August 3, 1993

APPLICATION NUMBER: 321344.5

FILING DATE: October 15, 1993

ATTORNEY/AGENT INFORMATION:

NAME: No. 6207814man D. Hanson

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5298.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 318-3000

TELEFAX: (212) 752-5958

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 502 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-382-256-12

Query Match 2.0%; Score 9; DB 4; Length 502;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
| | | | |  
Db 327 HRDLKSRNV 335

RESULT 19  
US-09-395-115-12  
Sequence 12, Application US/09395115  
Patent No. 6271365

GENERAL INFORMATION:  
APPLICANT: Miyazono, Kohei; DiJke, Peter Ten;  
APPLICANT: Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik  
TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins

TITLE OF INVENTION: Having Serine Threonine Kinase Domains And Their Use  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/395,115  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,265  
FILING DATE: 30-October-1995  
APPLICATION NUMBER: PCT/GB93/02367  
FILING DATE: 17-No. 6271365ember-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9224057.1  
FILING DATE: 17-No. 6271365ember-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9304677.9  
FILING DATE: 8-March-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9304680.3  
FILING DATE: 8-March-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9311047.6  
FILING DATE: 28-May-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9313763.6  
FILING DATE: 2-July-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9136099.2  
FILING DATE: 3-August-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9321344.5  
FILING DATE: 15-October-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohlei, Vineet  
REGISTRATION NUMBER: 37,003  
REFERENCE/DOCKET NUMBER: LUD 5298  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-395-115-12

Query Match 2.0%; Score 9; DB 4; Length 502;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
|||  
Db 327 HRDLKSRNV 335

## RESULT 20

US-08-436-265-12  
; Sequence 12, Application US/08436265  
; Patent No. 6316217  
; GENERAL INFORMATION:  
; APPLICANT: Miyazono, Kohlei; Dijke, Peter Ten;

APPLICANT: Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik  
TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,265  
FILING DATE: 30-October-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB93/02367  
FILING DATE: 17-No. 6316217ember-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9224057.1  
FILING DATE: 17-No. 6316217ember-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9304677.9  
FILING DATE: 8-March-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9304680.3  
FILING DATE: 8-March-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9311047.6  
FILING DATE: 28-May-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9313763.6  
FILING DATE: 2-July-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9136099.2  
FILING DATE: 3-August-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9321344.5  
FILING DATE: 15-October-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohlei, Vineet  
REGISTRATION NUMBER: 37,003  
REFERENCE/DOCKET NUMBER: LUD 5298  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-436-265-12

Query Match 2.0%; Score 9; DB 4; Length 502;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
|||  
Db 327 HRDLKSRNV 335

## RESULT 21

US-09-679-187-12  
; Sequence 12, Application US/09679187  
; Patent No. 6331621  
; GENERAL INFORMATION:  
; APPLICANT: Miyazono, Kohlei; Dijke, Peter Ten;

APPLICANT: Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik  
TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins  
TITLE OF INVENTION: Having Serine Threonine Kinase Domains And Their Use  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/679,187  
FILING DATE: 03-OCT-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,265  
FILING DATE: 30-October-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB93/02367  
FILING DATE: 17-NO. 6331621ember-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9224057.1  
FILING DATE: 17-NO. 6331621ember-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9304677.9  
FILING DATE: 8-March-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9304680.3  
FILING DATE: 8-March-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9311047.6  
FILING DATE: 28-May-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9313763.6  
FILING DATE: 2-July-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9136099.2  
FILING DATE: 3-August-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9321344.5  
FILING DATE: 15-October-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohlei, Vineet  
REGISTRATION NUMBER: 37,003  
REFERENCE/DOCKET NUMBER: LUD 5298  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-679-187-12

Query Match 2.0%; Score 9; DB 4; Length 502;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
| | | | | | | | | |  
Db 327 HRDLKSRNV 335

RESULT 22  
US-08-149-105-16  
; Sequence 16, Application US/08149105  
; Patent No. 5538892

GENERAL INFORMATION:  
APPLICANT: Donahoe, Patricia K.  
APPLICANT: Gustafson, Michael  
APPLICANT: He, Wei W.  
APPLICANT: Wang, Xiao-Fan  
TITLE OF INVENTION: TGF- TYPE I RECEPTOR  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: Wordperfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/149,105  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/029,673  
FILING DATE: March 11, 1993  
APPLICATION NUMBER: 07/853,396  
FILING DATE: March 18, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/211001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 505  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-149-105-16

Query Match 2.0%; Score 9; DB 1; Length 505;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
| | | | | | | | | |  
Db 330 HRDLKSRNV 338

RESULT 23  
US-08-317-847-16  
; Sequence 16, Application US/08317847  
; Patent No. 5547854  
; GENERAL INFORMATION:  
APPLICANT: Donahoe, Patricia K.  
APPLICANT: Gustafson, Michael  
APPLICANT: He, Wei W.  
TITLE OF INVENTION: FOUR NOVEL RECEPTORS OF THE TGF-B  
TITLE OF INVENTION: FAMILY  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/317,847  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/029,673  
FILING DATE: March 11, 1993  
APPLICATION NUMBER: 07/853,396  
FILING DATE: March 18, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/127002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 505  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-317-847-16

Query Match 2.0%; Score 9; DB 1; Length 505;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
|||  
Db 330 HRDLKSRNV 338

RESULT 24  
US-08-446-855A-2  
Sequence 2, Application US/08446855A  
Patent No. 5849573  
GENERAL INFORMATION:  
APPLICANT: Stewart, Thomas S  
APPLICANT: Flores, Maria V  
APPLICANT: O'Sullivan, William J  
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl  
TITLE OF INVENTION: phosphate synthetase II  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye PC  
STREET: 1100 No. 5849573th Glebe Road, 8th floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,855A  
FILING DATE: 06-Jul-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mitchell, Leonard C  
REGISTRATION NUMBER: 29,009  
REFERENCE/DOCKET NUMBER: 47-80  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:

LENGTH: 2391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-446-855A-2

Query Match 2.0%; Score 9; DB 2; Length 2391;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGE 445  
|||  
Db 1786 DDDDDDDGE 1794

RESULT 25  
US-09-150-741-2  
Sequence 2, Application US/09150741  
Patent No. 6183996  
GENERAL INFORMATION:  
APPLICANT: Stewart et al.  
TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate  
TITLE OF INVENTION: Synthetase II  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/09/150,741  
CURRENT FILING DATE: 1998-09-10  
EARLIER APPLICATION NUMBER: PL6380  
EARLIER FILING DATE: 1992-12-16  
EARLIER APPLICATION NUMBER: AU93/00617  
EARLIER FILING DATE: 1993-12-02  
EARLIER APPLICATION NUMBER: 08/446,855  
EARLIER FILING DATE: 1995-07-06  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 2391  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: protein  
US-09-150-741-2

Query Match 2.0%; Score 9; DB 4; Length 2391;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGE 445  
|||  
Db 1786 DDDDDDDGE 1794

RESULT 26  
US-08-240-712-35  
Sequence 35, Application US/08240712  
Patent No. 5599907  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, DAVID C.  
APPLICANT: MATHEWS, ANTONY JAMES  
APPLICANT: STETLER, GARY L.  
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC  
TITLE OF INVENTION: HEMOGLOBINS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Browdy and Nelmark  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/240,712  
FILING DATE: 09-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/09752  
FILING DATE: 13-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: ANDERSON=6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
OTHER INFORMATION: /note= one or both of Gly  
OTHER INFORMATION: residues 2 and 3 can be absent; any or all  
OTHER INFORMATION: of Asp residues 5-33 can be absent; one  
OTHER INFORMATION: or both of Gly residues 35 and 36 can be  
OTHER INFORMATION: absent  
MOLECULE TYPE: peptide  
US-08-240-712-35

Query Match 1.8%; Score 8; DB 1; Length 36;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDDDD 443  
|||  
Db 3 GDDDDDDDD 10

RESULT 27  
US-08-443-890-35  
Sequence 35, Application US/08443890  
Patent No. 5739011  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, DAVID C.  
APPLICANT: MATHEWS, ANTONY JAMES  
APPLICANT: STETLER, GARY L.  
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Browdy and Neimark  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/443,890  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/240,712  
FILING DATE: 09-MAY-1994  
APPLICATION NUMBER: PCT/US92/09752

FILING DATE: 13-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: ANDERSON=6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
OTHER INFORMATION: /note= one or both of Gly  
OTHER INFORMATION: residues 2 and 3 can be absent; any or all  
OTHER INFORMATION: of Asp residues 5-33 can be absent; one  
OTHER INFORMATION: or both of Gly residues 35 and 36 can be  
OTHER INFORMATION: absent  
MOLECULE TYPE: peptide  
US-08-443-890-35

Query Match 1.8%; Score 8; DB 1; Length 36;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDDDD 443  
|||  
Db 3 GDDDDDDDD 10

RESULT 28  
US-09-303-814-1  
Sequence 1, Application US/09303814  
Patent No. 6391848  
GENERAL INFORMATION:  
APPLICANT: de Lumen, Benito O.  
APPLICANT: Galvez, Alfredo F.  
TITLE OF INVENTION: Soybean Protein Nutraceuticals  
FILE REFERENCE: B99-089  
CURRENT APPLICATION NUMBER: US/09/303,814  
CURRENT FILING DATE: 1999-04-30  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 1  
LENGTH: 43  
TYPE: PRT  
ORGANISM: soybean  
US-09-303-814-1

Query Match 1.8%; Score 8; DB 4; Length 43;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDDDD 443  
|||  
Db 34 GDDDDDDDD 41

RESULT 29  
US-08-618-911-4  
Sequence 4, Application US/08618911  
Patent No. 5850016  
GENERAL INFORMATION:  
APPLICANT: Jung, Rudolf  
APPLICANT: Hastings, Craig  
APPLICANT: Coughlan, Sean  
APPLICANT: Hu, David  
TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN  
NUMBER OF SEQUENCES: 13

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.  
;; STREET: 700 Capital Square, 400 Locust Street  
;; CITY: Des Moines  
;; STATE: Iowa  
;; COUNTRY: USA  
;; ZIP: 50309  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/618,911  
;; FILING DATE: Concurrently herewith  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Simon, Soma  
;; REGISTRATION NUMBER: 37,444  
;; REFERENCE/DOCKET NUMBER: 365-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (515) 248-4800  
;; TELEFAX: (515) 248-4844  
;;  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 158 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-618-911-4

Query Match 1.8%; Score 8; DB 2; Length 158;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
Db 55 GDDDDDD 62

RESULT 30  
US-08-618-911-6  
; Sequence 6, Application US/08618911  
; Patent No. 5850016  
; GENERAL INFORMATION:  
; APPLICANT: Jung, Rudolf  
; APPLICANT: Haslings, Craig  
; APPLICANT: Coughlan, Sean  
; APPLICANT: Hu, David  
; TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN  
; TITLE OF INVENTION: SEEDS  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.  
; STREET: 700 Capital Square, 400 Locust Street  
; CITY: Des Moines  
; STATE: Iowa  
; COUNTRY: USA  
; ZIP: 50309  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/618,911  
; FILING DATE: Concurrently herewith  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Simon, Soma  
; REGISTRATION NUMBER: 37,444  
; REFERENCE/DOCKET NUMBER: 365-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (515) 248-4800  
; TELEFAX: (515) 248-4844

;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 158 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-618-911-6

Query Match 1.8%; Score 8; DB 2; Length 158;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
Db 55 GDDDDDD 62

RESULT 31  
US-08-938-675A-2  
; Sequence 2, Application US/08938675A  
; Patent No. 6107287  
; GENERAL INFORMATION:  
; APPLICANT: de Lumen, Benito O.  
; APPLICANT: Galvez, Alfredo F.  
; TITLE OF INVENTION: Lunasin Peptides  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 75 DENISE DRIVE  
; CITY: HILLSBOROUGH  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94010  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/938,675A  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B98-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 343-4341  
; TELEFAX: (650) 343-4342  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 158 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-938-675A-2

Query Match 1.8%; Score 8; DB 3; Length 158;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
Db 55 GDDDDDD 62

RESULT 32  
US-08-752-132-2  
; Sequence 2, Application US/08752132  
; Patent No. 5869290  
; GENERAL INFORMATION:  
; APPLICANT: Freeman, Kathryn

APPLICANT: Nicholas, Richard  
APPLICANT: Livi, George  
TITLE OF INVENTION: CAYAEI  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/752,132  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50580  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 179 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
US-08-752-132-2

Query Match 1.8%; Score 8; DB 2; Length 179;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 436 GDDDDDDDD 443  
Db 117 GDDDDDDDD 124

RESULT 33  
US-09-382-080-1  
Sequence 1, Application US/09382080  
Patent No. 6087333  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: DISEASE ASSOCIATED ACIDIC PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/382,080  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/859,937  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0308 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: THPINOT03  
CLONE: 2448450  
US-09-382-080-1

Query Match 1.8%; Score 8; DB 3; Length 192;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 436 GDDDDDDDD 443  
Db 118 GDDDDDDDD 125

RESULT 34  
US-08-859-937-1  
Sequence 1, Application US/08859937  
Patent No. 6090577  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: DISEASE ASSOCIATED ACIDIC PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/859,937  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0308 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: THPINOT03  
CLONE: 2448450  
US-08-859-937-1

Query Match 1.8%; Score 8; DB 3; Length 192;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
Db 118 GDDDDDD 125

RESULT 35  
US-08-431-080-26  
Sequence 26, Application US/08431080  
Patent No. 5698686  
GENERAL INFORMATION:

APPLICANT: Gottschling, Daniel E.  
Singer, Miriam S.  
TITLE OF INVENTION: Telomerase Compositions and Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TEXAS  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,080  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SN 08/326,781  
FILING DATE: October 20, 1994

CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 226 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-431-080-26

Query Match 1.8%; Score 8; DB 1; Length 226;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 444  
Db 142 DDDDDDD 149

RESULT 36  
US-08-938-534-26  
Sequence 26, Application US/08938534  
Patent No. 5916752

## GENERAL INFORMATION:

APPLICANT: Gottschling, Daniel E.  
Singer, Miriam S.  
TITLE OF INVENTION: Telomerase Compositions and Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TEXAS  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/938,534  
FILING DATE: 26-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/431,080  
FILING DATE:  
APPLICATION NUMBER: SN 08/326,781  
FILING DATE: October 20, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 226 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-938-534-26

Query Match 1.8%; Score 8; DB 2; Length 226;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 444  
Db 142 DDDDDDD 149

RESULT 37  
US-09-345-294-26  
Sequence 26, Application US/09345294  
Patent No. 6387619  
GENERAL INFORMATION:

APPLICANT: Gottschling, Daniel E.  
Singer, Miriam S.  
TITLE OF INVENTION: Telomerase Compositions and Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TEXAS  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:



APPLICATION NUMBER: US/09/345,294  
FILING DATE: 30-JUN-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/431,080  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 226 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
\* TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
-09-345-294-26

Query Match 1.8%; Score 8; DB 4; length 226;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
Db 142 DDDDDDDG 149

RESULT 38  
US-08-469-537A-80  
Sequence 80, Application US/08469537A  
Patent No. 5843749  
GENERAL INFORMATION:  
APPLICANT: Maisongier, et al.  
TITLE OF INVENTION: EHK AND ROR TYROSINE  
TITLE OF INVENTION: KINASES  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: NY  
COUNTRY: U.S.A.  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,537A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/406,247  
FILING DATE: 17-MAR-1995  
APPLICATION NUMBER: USSN 08/144,992  
FILING DATE: 28-OCT-1993  
APPLICATION NUMBER: USSN 07/736,559  
FILING DATE: 26-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kempler, Ph.D., Gail M  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: REG 070C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
TELEX:  
INFORMATION.FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-469-537A-80

Query Match 1.8%; Score 8; DB 2; length 251;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVLWE 198  
Db 109 SYGVVLWE 116

RESULT 39  
US-08-469-537A-81  
Sequence 81, Application US/08469537A  
Patent No. 5843749  
GENERAL INFORMATION:  
APPLICANT: Maisongier, et al.  
TITLE OF INVENTION: EHK AND ROR TYROSINE  
TITLE OF INVENTION: KINASES  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: NY  
COUNTRY: U.S.A.  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,537A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/406,247  
FILING DATE: 17-MAR-1995  
APPLICATION NUMBER: USSN 08/144,992  
FILING DATE: 28-OCT-1993  
APPLICATION NUMBER: USSN 07/736,559  
FILING DATE: 26-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kempler, Ph.D., Gail M  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: REG 070C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
TELEX:  
INFORMATION FOR SEQ ID NO: 81:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-469-537A-81

Query Match 1.8%; Score 8; DB 2; length 251;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVLWE 198  
Db 109 SYGVVLWE 116

```
RESULT 40
US-08-852-743-5
; Sequence 5, Application US/08852743
; Patent No. 5830699
; GENERAL INFORMATION:
; APPLICANT: Force, Thomas
; APPLICANT: Kyriakis, John M.
; APPLICANT: Pombo, Celia M.
; APPLICANT: Bonventre, Joseph
; TITLE OF INVENTION: SOK-1 AND METHODS OF USE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,743
; FILING DATE: 7-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,774
; FILING DATE: 7-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/327001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-852-743-5

Query Match      1.8%; Score 8; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
Db 165 WMAPEVIQ 172

RESULT 41
US-09-185-370-5
; Sequence 5, Application US/09185370
; Patent No. 6093560
; GENERAL INFORMATION:
; APPLICANT: Force, Thomas
; APPLICANT: Kyriakis, John M.
; APPLICANT: Pombo, Celia M.
; APPLICANT: Bonventre, Joseph
; TITLE OF INVENTION: SOK-1 AND METHODS OF USE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
```

```
ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/185,370
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/852,743
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/327001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-185-370-5
```

```
Query Match      1.8%; Score 8; DB 3; Length 270;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
Db 165 WMAPEVIQ 172
```

```
RESULT 42
US-09-485-077A-17
; Sequence 17, Application US/09485077A
; Patent No. 6458590
; GENERAL INFORMATION:
; APPLICANT: Mukherjee, Anil
; APPLICANT: Kundu, Gopal
; APPLICANT: Panda, Dibyendu
; TITLE OF INVENTION: Methods and Compositions for Treatment of Restenosis
; FILE REFERENCE: NIH-05047
; CURRENT APPLICATION NUMBER: US/09/485,077A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: PCT/US98/16569
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: 60/054,967
; PRIOR FILING DATE: 1997-07-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Rattus No. 6458590vegicus
; US-09-485-077A-17
```

```
Query Match      1.8%; Score 8; DB 4; Length 317;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
Db 88 DDDDDDDG 95
```

```
RESULT 43
US-09-211-930-3
```

```
; Sequence 3, Application US/09211930
; Patent No. 5962265
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 5962265ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
; TITLE OF INVENTION: SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296
; CURRENT APPLICATION NUMBER: US/09/211,930
; CURRENT FILING DATE: 1998-12-15
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 3
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-211-930-3
```

```
Query Match 1.8%; Score 8; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 172 WMAPEVIQ 179
Db 185 WMAPEVIQ 192
```

```
RESULT 44
US-09-211-930-11
; Sequence 11, Application US/09211930
; Patent No. 5962265
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 5962265ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
; TITLE OF INVENTION: SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296
; CURRENT APPLICATION NUMBER: US/09/211,930
; CURRENT FILING DATE: 1998-12-15
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 11
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-211-930-11
```

```
Query Match 1.8%; Score 8; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 172 WMAPEVIQ 179
Db 185 WMAPEVIQ 192
```

```
RESULT 45
US-09-340-993-3
; Sequence 3, Application US/09340993
; Patent No. 6034228
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6034228ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.N1
; CURRENT APPLICATION NUMBER: US/09/340,993
```

```
; CURRENT FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0 & US 09/211,930
; EARLIER FILING DATE: 1997-12-19 & 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 3
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-340-993-3
```

```
Query Match 1.8%; Score 8; DB 3; Length 416;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 172 WMAPEVIQ 179
Db 185 WMAPEVIQ 192
```

```
RESULT 46
US-09-340-993-11
; Sequence 11, Application US/09340993
; Patent No. 6034228
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6034228ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.N1
; CURRENT APPLICATION NUMBER: US/09/340,993
; CURRENT FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0 & US 09/211,930
; EARLIER FILING DATE: 1997-12-19 & 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 11
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-993-11
```

```
Query Match 1.8%; Score 8; DB 3; Length 416;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 172 WMAPEVIQ 179
Db 185 WMAPEVIQ 192
```

```
RESULT 47
US-09-468-442-3
; Sequence 3, Application US/09468442
; Patent No. 6300098
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6300098ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.N1
; CURRENT APPLICATION NUMBER: US/09/468,442
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/340,993
; EARLIER FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; EARLIER APPLICATION NUMBER: US 09/211,930
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 3
; LENGTH: 416
```

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-468-442-3

Query Match 1.8%; Score 8; DB 4; Length 416;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
|||||  
Db 185 WMAPEVIQ 192

## RESULT 48

US-09-468-442-11  
Sequence 11, Application US/09468442  
Patent No. 6300098  
GENERAL INFORMATION:  
APPLICANT: Tyrell E. No. 6300098r18  
APPLICANT: William Craig Moore  
APPLICANT: David Shay Silberstein  
TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE  
FILE REFERENCE: PHM 70296.N1  
CURRENT APPLICATION NUMBER: US/09/468,442  
CURRENT FILING DATE: 1999-12-21  
EARLIER APPLICATION NUMBER: US 09/340,993  
EARLIER FILING DATE: 1999-06-25  
EARLIER APPLICATION NUMBER: GB 9726851.0  
EARLIER FILING DATE: 1997-12-19  
EARLIER APPLICATION NUMBER: US 09/211,930  
EARLIER FILING DATE: 1998-12-15  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 11  
LENGTH: 416  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-468-442-11

Query Match 1.8%; Score 8; DB 4; Length 416;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
|||||  
Db 185 WMAPEVIQ 192

## RESULT 49

US-08-712-709-8  
Sequence 8, Application US/08712709  
Patent No. 5863780  
GENERAL INFORMATION:  
APPLICANT: Au-Young, Janice  
APPLICANT: Guegler, Karl J.  
APPLICANT: Hawkins, Phillip R.  
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/712,709  
FILING DATE: Filed Herewith

ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0118 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1117791  
US-08-712-709-8

Query Match 1.8%; Score 8; DB 2; Length 487;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
|||||  
Db 190 WMAPEVIQ 197

## RESULT 50

US-09-111-444-8  
Sequence 8, Application US/09111444  
Patent No. 6045792  
GENERAL INFORMATION:  
APPLICANT: Au-Young, Janice  
APPLICANT: Guegler, Karl J.  
APPLICANT: Hawkins, Phillip R.  
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/111,444  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/712,709  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0118 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1117791



US-09-111-444-8

Query Match 1.8%; Score 8; DB 3; Length 487;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
|||||  
Db 190 WMAPEVIQ 197

RESULT 51

US-09-541-228-8

; Sequence 8, Application US/09541228  
; Patent No. 6232077  
; GENERAL INFORMATION:

APPLICANT: Au-Young, Janice  
APPLICANT: Guegler, Karl J.  
APPLICANT: Hawkins, Phillip R.  
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/541,228  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/712,709  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0118 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1117791

US-09-541-228-8

Query Match 1.8%; Score 8; DB 4; Length 487;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
|||||  
Db 190 WMAPEVIQ 197

RESULT 52

US-09-529-279-4

; Sequence 4, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:  
; ADDRESSEE: ONO, KOICHIRO  
; APPLICANT: ONO, KOICHIRO

APPLICANT: OHTOMO, TOSHIHIKO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/09/529,279  
CURRENT FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188  
PRIOR FILING DATE: 1997-10-22  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 579  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-529-279-4

Query Match 1.8%; Score 8; DB 4; Length 579;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
|||||  
Db 170 VLKICDFG 177

RESULT 53

US-09-529-279-15

; Sequence 15, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:

APPLICANT: ONO, KOICHIRO  
APPLICANT: OHTOMO, TOSHIHIKO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/09/529,279  
CURRENT FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188  
PRIOR FILING DATE: 1997-10-22  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 15  
LENGTH: 590  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-529-279-15

Query Match 1.8%; Score 8; DB 4; Length 590;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
|||||  
Db 170 VLKICDFG 177

RESULT 54

US-08-469-537A-107

; Sequence 107, Application US/08469537A  
; Patent No. 5843749  
; GENERAL INFORMATION:  
APPLICANT: Maisonnier, et al.  
TITLE OF INVENTION: EHK AND ROR TYROSINE  
KINASES  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown

STATE: NY  
COUNTRY: U.S.A.  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,537A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/406,247  
FILING DATE: 17-MAR-1995  
APPLICATION NUMBER: USSN 08/144,992  
FILING DATE: 28-OCT-1993  
APPLICATION NUMBER: USSN 07/736,559  
FILING DATE: 26-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kempler, Ph.D., Gall M  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: REG 070C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
TELEX:  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 943 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Human ROR2  
LOCATION: 1..943  
OTHER INFORMATION:  
US-08-469-537A-107

Query Match 1.8%; Score 8; DB 2; Length 943;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVLWE 198  
|||||||  
677 SYGVVLWE 684

RESULT 55  
US-08-431-080-28  
Sequence 28, Application US/08431080  
Patent No. 5698686  
GENERAL INFORMATION:  
APPLICANT: Gottschling, Daniel E.  
APPLICANT: Singer, Miriam S.  
TITLE OF INVENTION: Telomerase Compositions and Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TEXAS  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,080

FILING DATE: Concurrently Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SN 08/326,781  
FILING DATE: October 20, 1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1085 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-431-080-28

Query Match 1.8%; Score 8; DB 1; Length 1085;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
|||||||  
Db 142 DDDDDDDG 149

RESULT 56  
US-08-938-534-28  
Sequence 28, Application US/08938534  
Patent No. 5916752  
GENERAL INFORMATION:  
APPLICANT: Gottschling, Daniel E.  
APPLICANT: Singer, Miriam S.  
TITLE OF INVENTION: Telomerase Compositions and Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TEXAS  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/938,534  
FILING DATE: 26-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/431,080  
FILING DATE:  
APPLICATION NUMBER: SN 08/326,781  
FILING DATE: October 20, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1085 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-938-534-28

Query Match 1.8%; Score 8; DB 2; Length 1085;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
Db 142 DDDDDDDG 149

RESULT 57  
US-09-345-294-28  
Sequence 28, Application US/09345294  
Patent No. 6387619  
GENERAL INFORMATION:  
APPLICANT: Gottschling, Daniel E.

TITLE OF INVENTION: Telomerase Compositions and Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TEXAS  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/345,294  
FILING DATE: 30-Jun-1999  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/431,080  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1085 amino acids  
\* TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-09-345-294-28

Query Match 1.8%; Score 8; DB 4; Length 1085;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
Db 142 DDDDDDDG 149

RESULT 58  
US-08-482-228-215  
Sequence 215, Application US/08482228  
Patent No. 5968753  
GENERAL INFORMATION:  
APPLICANT: Tseng-Law, Janet

APPLICANT: Kobori, Joan A.  
APPLICANT: Al-Abdaly, Fahad A.  
APPLICANT: Guillermo, Roy  
APPLICANT: Helgeson, Sam L.  
APPLICANT: Deans, Robert J.  
TITLE OF INVENTION: POSITIVE AND NEGATIVE CELL  
TITLE OF INVENTION: SELECTION MEDIATED BY PEPTIDE RELEASE  
NUMBER OF SEQUENCES: 215  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Janice Guthrie, Ph.D.  
STREET: P.O. Box 15210  
CITY: Irvine  
STATE: California  
COUNTRY: USA  
ZIP: 92713-5210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,228  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Guthrie, Janice  
REGISTRATION NUMBER: 35,170  
REFERENCE/DOCKET NUMBER: IT-4630CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (714) 440-5353  
TELEFAX: (714) 553-1952

INFORMATION FOR SEQ ID NO: 215:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-482-228-215

Query Match 1.5%; Score 7; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443  
Db 1 DDDDDDD 7

RESULT 59  
US-08-482-528-215  
Sequence 215, Application US/08482528  
Patent No. 6017719  
GENERAL INFORMATION:

APPLICANT: Tseng-Law, Janet  
APPLICANT: Kobori, Joan A.  
APPLICANT: Al-Abdaly, Fahad A.  
APPLICANT: Helgeson, Roy  
APPLICANT: Deans, Robert J.  
TITLE OF INVENTION: POSITIVE AND NEGATIVE CELL  
TITLE OF INVENTION: SELECTION MEDIATED BY PEPTIDE RELEASE  
NUMBER OF SEQUENCES: 215  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Janice Guthrie, Ph.D.  
STREET: P.O. Box 15210  
CITY: Irvine  
STATE: California  
COUNTRY: USA  
ZIP: 92713-5210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,528  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Guthrie, Janice  
REGISTRATION NUMBER: 35,170  
REFERENCE/DOCKET NUMBER: IT-4630CIP4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (714) 440-5353  
TELEFAX: (714) 553-1952  
INFORMATION FOR SEQ ID NO: 215:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-482-528-215

Query Match 1.5%; Score 7; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 1 DDDDDDD 7

RESULT 60  
US-08-199-776-23  
Sequence 23, Application US/08199776  
Patent No. 5594120  
GENERAL INFORMATION:  
APPLICANT: Brenner, Michael B.  
APPLICANT: Parker, Christina M.  
TITLE OF INVENTION: No. 5594120el integrin alpha subunit  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/199,776  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: synthetic peptide  
US-08-199-776-23

Query Match 1.5%; Score 7; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 9.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 3 DDDDDDD 9

RESULT 61  
US-08-663-731-23  
Sequence 23, Application US/08663731  
Patent No. 6057423  
GENERAL INFORMATION:  
APPLICANT: Brenner, Michael B.  
APPLICANT: Parker, Christina M.  
TITLE OF INVENTION: No. 6057423el integrin alpha subunit  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/663,731  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/199,776  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: synthetic peptide  
US-08-663-731-23

Query Match 1.5%; Score 7; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 9.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 3 DDDDDDD 9

RESULT 62  
US-08-879-338-23  
Sequence 23, Application US/08879338A  
Patent No. 6063906  
GENERAL INFORMATION:  
APPLICANT: Brenner, Michael B.



```

; APPLICANT: Parker, Christina M.
; TITLE OF INVENTION: Antibodies to No. 6063906el Integrin Alpha
; FILE REFERENCE: B0801/7080/ERP
; CURRENT APPLICATION NUMBER: US/08/879,338A
; CURRENT FILING DATE: 1997-06-20
; EARLIER APPLICATION NUMBER: US 08/663,731
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: US 08/199,776
; EARLIER FILING DATE: 1994-02-18
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
; US-08-879-338-23

Query Match
Best Local Similarity 1.5%; Score 7; DB 3; Length 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
   |||||
Db 3 DDDDDDD 9

RESULT 63
US-09-293-238B-23
; Sequence 23, Application US/09293238B
; Patent No. 6455042
; GENERAL INFORMATION:
; APPLICANT: Brenner, Michael B.
; APPLICANT: Parker, Christina M.
; TITLE OF INVENTION: A Method of Treating Ulcerative Colitis
; TITLE OF INVENTION: or Crohn's Disease by Administering an Antibody to Alpha E
; TITLE OF INVENTION: Beta 7 Integrin
; FILE REFERENCE: L0560/7005/ERP
; CURRENT APPLICATION NUMBER: US/09/293,238B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: US 08/879,338
; PRIOR FILING DATE: 1997-06-20
; PRIOR APPLICATION NUMBER: US 08/663,731
; PRIOR FILING DATE: 1996-06-14
; PRIOR APPLICATION NUMBER: US 08/199,776
; PRIOR FILING DATE: 1994-02-18
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
; US-09-293-238B-23

Query Match
Best Local Similarity 1.5%; Score 7; DB 4; Length 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
   |||||
Db 3 DDDDDDD 9

RESULT 64
PCT-US95-02044-23
; Sequence 23, Application PC/TUS9502044
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Novel integrin alpha subunit
```

```

; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02044
; FILING DATE: herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/199,776
; FILING DATE: 18 February 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: synthetic peptide
; PCT-US95-02044-23

Query Match
Best Local Similarity 1.5%; Score 7; DB 5; Length 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
   |||||
Db 3 DDDDDDD 9

RESULT 65
US-09-625-570-4
; Sequence 4, Application US/09625570
; Patent No. 6437095
; GENERAL INFORMATION:
; APPLICANT: Lillie, Hauke
; APPLICANT: Richter, Susanne
; APPLICANT: Rudolph, Rainer
; APPLICANT: Stubenrauch, Kay-Gunnar
; TITLE OF INVENTION: Chimeric polypeptides, method for production and uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: Case 20393
; CURRENT APPLICATION NUMBER: US/09/625,570
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: EP99115022.8
; PRIOR FILING DATE: 1999-08-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide D10SCP
```

US-09-625-570-4

Query Match 1.5%; Score 7; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
|||||  
Db 1 DDDDDDD 7

RESULT 66

US-09-039-780A-80  
; Sequence 80, Application US/09039780A  
; Patent No. 6376248

GENERAL INFORMATION:

APPLICANT: HAWLEY-NELSON, PAMELA  
LAN, JIANQING  
SHIH, POJEN  
JESSE, JOEL A.  
SCHIFFERLI, KEVIN P.  
GEBEYEHU, GULILAT

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS  
NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: GREENLEE, WINNER & SULLIVAN  
STREET: 5370 MANHATTAN CIRCLE, SUITE 201  
CITY: BOULDER  
STATE: CO  
COUNTRY: US  
ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/039,780A  
FILING DATE: 16-Mar-1998  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: SULLIVAN, SALLY A.  
REGISTRATION NUMBER: 32,064  
REFERENCE/DOCKET NUMBER: 32-95C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303)499-8080  
TELEFAX: (303)499-8089

INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: Modified-site

LOCATION: 14

OTHER INFORMATION: /product= "OTHER"

/note= "G AT POSITION 14 CAN BE ABSENT"

SEQUENCE DESCRIPTION: SEQ ID NO: 80:

US-09-039-780A-80

Query Match 1.5%; Score 7; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 DDDDDDG 444  
|||||  
Db 8 DDDDDDG 14

RESULT 67

US-09-039-780A-87

; Sequence 87, Application US/09039780A  
; Patent No. 6376248

GENERAL INFORMATION:

APPLICANT: HAWLEY-NELSON, PAMELA  
LAN, JIANQING  
SHIH, POJEN  
JESSE, JOEL A.  
SCHIFFERLI, KEVIN P.  
GEBEYEHU, GULILAT

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS  
NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: GREENLEE, WINNER & SULLIVAN  
STREET: 5370 MANHATTAN CIRCLE, SUITE 201  
CITY: BOULDER  
STATE: CO  
COUNTRY: US  
ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/039,780A  
FILING DATE: 16-Mar-1998  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: SULLIVAN, SALLY A.  
REGISTRATION NUMBER: 32,064  
REFERENCE/DOCKET NUMBER: 32-95C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303)499-8080  
TELEFAX: (303)499-8089

INFORMATION FOR SEQ ID NO: 87:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: Modified-site

LOCATION: 30

OTHER INFORMATION: /product= "OTHER"

/note= "G AT POSITION 30 CAN BE ABSENT"

SEQUENCE DESCRIPTION: SEQ ID NO: 87:

US-09-039-780A-87

Query Match 1.5%; Score 7; DB 4; Length 30;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 DDDDDDG 444  
|||||  
Db 24 DDDDDDG 30

RESULT 68

US-09-376-113-3

; Sequence 3, Application US/09376113  
; Patent No. 6451992

GENERAL INFORMATION:

APPLICANT: Cupp, Eddie Wayne  
APPLICANT: Cupp, Mary Smith  
TITLE OF INVENTION: Antithrombin Nucleotides and Proteins  
TITLE OF INVENTION: from Horn Fly  
FILE REFERENCE: 5721-10  
CURRENT APPLICATION NUMBER: US/09/376,113

CURRENT FILING DATE: 1999-08-17  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 30  
TYPE: PRT  
ORGANISM: Haematobia irritans  
US-09-376-113-3

Query Match 1.5%; Score 7; DB 4; Length 30;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443  
Db 10 DDDDDDD 16

RESULT 69

US-08-290-736C-9  
Sequence 9, Application US/08290736C  
Patent No. 6294174

GENERAL INFORMATION:

APPLICANT: KRSMANOVIC, VELIBOR

COSIC, IRENA

BIOUARD, JEAN-MICHEL

HEARN, MILTON TW

TITLE OF INVENTION: PEPTIDES IMMUNOLOGICALLY RELATED TO

PROTEINS OF A VIRAL AGENT AND THEIR BIOLOGICAL APPLICATION

NUMBER OF SEQUENCES: 48

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHAYE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VA

COUNTRY: USA

ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/290,736C

FILING DATE: 16-NO. 6294174-1994

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/FR93/00171

FILING DATE: 19-FEB-1993

APPLICATION NUMBER: FR92/01883

FILING DATE: 19-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: SADOFF, B.J.

REGISTRATION NUMBER: 36663

REFERENCE/DOCKET NUMBER: 1721-3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 7038164000

TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 50 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..50

SEQUENCE DESCRIPTION: SEQ ID NO: 9:

US-08-290-736C-9

Query Match 1.5%; Score 7; DB 4; Length 50;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443  
Db 4 DDDDDDD 10

RESULT 70

US-08-290-736C-8  
Sequence 8, Application US/08290736C  
Patent No. 6294174

GENERAL INFORMATION:

APPLICANT: KRSMANOVIC, VELIBOR

COSIC, IRENA

BIOUARD, JEAN-MICHEL

HEARN, MILTON TW

TITLE OF INVENTION: PEPTIDES IMMUNOLOGICALLY RELATED TO

PROTEINS OF A VIRAL AGENT AND THEIR BIOLOGICAL APPLICATION

NUMBER OF SEQUENCES: 48

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHAYE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VA

COUNTRY: USA

ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/290,736C

FILING DATE: 16-NO. 6294174-1994

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/FR93/00171

FILING DATE: 19-FEB-1993

APPLICATION NUMBER: FR92/01883

FILING DATE: 19-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: SADOFF, B.J.

REGISTRATION NUMBER: 36663

REFERENCE/DOCKET NUMBER: 1721-3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 7038164000

TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 73 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..73

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-08-290-736C-8

Query Match 1.5%; Score 7; DB 4; Length 73;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443  
Db 40 DDDDDDD 46

RESULT 71

US-08-117-952-757

Sequence 757, Application US/08117952

Patent No. 5851760

GENERAL INFORMATION:

APPLICANT: Evans, Glen A.

```

; APPLICANT: Smith, Michael W.
; TITLE OF INVENTION: METHOD FOR GENERATION OF SEQUENCE
; TITLE OF INVENTION: SAMPLED MAPS OF COMPLEX GENOMES
; NUMBER OF SEQUENCES: 797
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/117,952
; FILING DATE: 07-SEP-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/078,471
; FILING DATE: 15-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9423
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4737
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 757:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-117-952-757

Query Match
Best Local Similarity 1.5%; Score 7; DB 2; Length 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRLK 135
DB 59 VIHRLK 65

RESULT 72
US-09-376-113-2
; Sequence 2, Application US/09376113
; Patent No. 6451992
; GENERAL INFORMATION:
; APPLICANT: Cupp, Eddie Wayne
; APPLICANT: Cupp, Mary Smith
; TITLE OF INVENTION: Antithrombin Nucleotides and Proteins
; FILE REFERENCE: 5721-10
; CURRENT APPLICATION NUMBER: US/09/376,113
; CURRENT FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Haematobia irritans
; US-09-376-113-2

Query Match
Best Local Similarity 1.5%; Score 7; DB 4; Length 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
```

```

DB 10 DDDDDDD 16

RESULT 73
US-08-370-225-20
; Sequence 20, Application US/08370225
; Patent No. 5580736
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; APPLICANT: Gyuris, Jeno
; APPLICANT: Golemis, Erica
; TITLE OF INVENTION: Interaction Trap System for Isolating
; TITLE OF INVENTION: No. 5580736e1 Proteins
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 555X
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,225
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/969,038
; FILING DATE: 10/30/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/143001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-370-225-20

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRLK 135
DB 47 VIHRLK 53

RESULT 74
US-08-461-859-20
; Sequence 20, Application US/08461859
; Patent No. 5786169
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; APPLICANT: Gyuris, Jeno
; APPLICANT: Golemis, Erica
; TITLE OF INVENTION: Interaction Trap System for Isolating
; TITLE OF INVENTION: No. 5786169e1 Proteins
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
```



CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: Wordperfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,859  
FILING DATE: June 5, 1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/370,225  
FILING DATE: January 9, 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/969,038  
FILING DATE: October 30, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lech, Karen F.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/143002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-461-859-20

Query Match 1.5%; Score 7; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRLK 135  
DB 47 VIHRLK 53

RESULT 75  
PCT-US93-10069-20  
Sequence 20, Application PC/TUS9310069  
GENERAL INFORMATION:  
APPLICANT: Brent, Roger  
APPLICANT: Gyuris, Jeno  
APPLICANT: Golemis, Erica  
TITLE OF INVENTION: Interaction Trap System for Isolating  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: Wordperfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/10069  
FILING DATE: 20-OCT-1993  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/969,038  
FILING DATE: 10/30/92

ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/143001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
PCT-US93-10069-20

Query Match 1.5%; Score 7; DB 5; Length 86;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRLK 135  
DB 47 VIHRLK 53

Search completed: May 1, 2003, 20:56:36  
Job time : 31 secs